Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE

Bc. Zuzana Valášková

Klasifikační tarifování

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Lucie Mazurová, Ph.D
Studijní program: Finanční a pojistná matematika

2008

Na tomto místě bych chtěla poděkovat všem, kteří mi pomohli při vypracování mé práce. V první řade moje děkuji patří vedoucímu mé diplomové práce RNDr. Lucií Mazurové, Ph.D., která mi vždy ochotně pomohla. Děkuji Knihovně Václava Hlavatého, rodině a všem, kteří mi či už přímo nebo nepřímo pomáhali.

Prohlašuji, že jsem svou diplomovou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 13.7.2009 Bc. Zuzana Valášková

Obsah

Název práce: Klasifikační tarifování
Autor: Bc. Zuzana Valášková
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí diplomové práce: RNDr. Lucie Mazurová, Ph.D
e-mail vedoucího: lucie.mazurova@mff.cuni.cz

Abstrakt: V tejto práci sa budeme zaoberať metódami aplikovateľnými pri určovaní poistného v neživotnom poistení na základe zaradenia do rizikovej skupiny pomocou určitých indicií. Celá práca je rozdelené do troch základných častí, pričom výsledky z každej tejto časti sú v závere ilustrované na príklade. V prvej časti sa budeme zaoberať základnými metódami tarifovania aplikovanými na aditívnu a multiplikatívnu tarifnú štruktúru. Za základné metódy budeme považovať váženú metódu najmenších štvorcov, metódu marginálnych súčtov a metódu Baileyho - Simonovu. V druhej časti sa budeme zaoberať sofistikovanejším prístupom k tarifovaniu a to prostredníctvom zobecnených lineárnych modelov. V tretej časti sa budeme zaoberať kredibilitnými modelmi v tarifovaní, ktoré majú v dnešnej dobe významné postavenie v modelovaní budúceho vývoja.

Klíčová slova: poistné, tarifná premenná, tarifná trieda, aditívna a multiplikatívna štruktúra

Title: The classification ratemaking
Author: Bc. Zuzana Valášková
Department: Department of Probability and Mathematical Statistics
Supervisor: RNDr. Lucie Mazurová, Ph.D
Supervisor's e-mail address: lucie.mazurova@mff.cuni.cz

Abstract: In the present work we will study methods, which are used to find a premium in nonlife insurance according to the grouping risks into the risk groups. These risk groups are constructed according to the concrete indices. All work consists of the three main sections. At the end you can find example, where all methods are applied. First section explains the basic methods which can be used in ratemaking: the weighted least square method, the method of marginal totals and the method of Bailey - Simon. The second section studies more sophisticated method for ratemaking: The generalized linear models. The third section studies the credibility estimations used in ratemaking, which are important and used today.

Keywords: premium, tariff variable, tariff class, additive and multiplicative structure.

Kapitola 1

Základná štruktúra

Cieľom klasifikačného tarifovania je odhadnúť budúci vývoj počtu a veľkosti škôd na základe znalosti **tarifnej triedy**. Pri tomto odhade sa využívajú dáta z minulých období. Tarifná trieda je homogénna skupina rizík (poistných zmluv), teda je možné vyžadovať jednotné poistné pre všetky riziká v rámci jednej tarifnej triedy. Riziká zaraďujeme do tarifných tried na základe znalosti hodnôt **tarifných premenných**. Tarifná premenná je parameter, ktorého hodnota nám, z určitého hľadiska, vypovedá o rizikovosti poisteného. Celý proces spočíva v tom, že na základe tarifných premenných rozdelíme riziká do tarifných tried, pre ktoré určíme poistné.

Príklad 1.
Uvažujeme poistenie motorového vozidla. Príklady tarifných premenných, ktoré nám dajú informáciu o rizikovosti poistenia môžu byť nasledovné.

1. tarifná premenná: **Vek vodiča**
Mladý vodič je vo všeobecnosti vodič neskúsený. Často prekračuje rýchlosť, nie vždy vie odhadnúť svoje možnosti. Z tohto pohľadu je pre nás rizikovejší ako vodič starší, ktorý väčšinou jazdí zodpovednejšie a všeobecne má viac skúseností.

2. tarifná premenná: **Pohlavie**
Muži sú vo všeobecnosti považovaný za rizikovejších vodičov. Ženy sú vo všeobecnosti opatrnejšie.

3. tarifná premenná: **Počet najazdených kilometrov za rok**
Týmto údajom získa poisťovňa predstavu o využívateľnosti motorového vozidla.

Takto by sme mohli vymenovať ďalšie tarifné premenné, napr. región, objem motoru, vek vozidla, počet najazdených rokov, údaje o priestupkoch vodiča v dôsledku používania motorového vozidla, zaobstarávajúcia cena vozidla a podobne.

V tomto texte budeme pre jednoduchosť a názornosť uvažovať dve tarifné premenné, ktoré si označíme A a B. Na základe hodnoty tarifnej premennej A zaradíme príslušné riziko do triedy z množiny $\{1, 2, 3, \ldots I\}$ a na základe hodnoty

Tabuľka 1.1: Klasifikačná štruktúra

I/J	**1**	**2**	**3**	**...**	**I**
1	$(1,1)$	$(1,2)$	$(1,3)$	...	$(1,J)$
2	$(2,1)$	$(2,2)$	$(2,3)$	...	$(2,J)$
3	$(3,1)$	$(3,2)$	$(3,3)$	...	$(3,J)$
⋮	...	...	(i,j)	...	(i,J)
J	$(I,1)$	...	...	...	(I,J)

tarifnej premennej B zaradíme riziko do triedy z množiny $\{1, 2, 3, \ldots J\}$. Týmto spôsobom zkonštruujeme $I \times J$ tarifných tried, čo vieme jednoducho znázorniť pomocou Tabuľky 1.1.

Nech X_{ij} je náhodná veličina sledovaná v tarifnej triede (i, j), a nech n_{ij} je miera vystavenia riziku tejto tarifnej triedy (miera expozície).

X_{ij} môže byť napríklad:

1. celkový objem škôd v tarifnej triede (i, j),
2. celkový počet škôd v tarifnej triede (i, j),
3. celkový objem škôd delený počtom zmluv v tarifnej triede (i, j),
4. priemerná veľkosť škody v tarifnej triede (i, j) a pod.

n_{ij} môže byť napríklad:

1. počet rokov,na ktoré je poistenie zjednané v tarifnej triede (i, j),
2. počet zmluv v tarifnej triede (i, j),
3. celkový počet škôd v tarifnej triede (i, j) a pod.

A čo bude našou úlohou?

Cieľom tejto práce je vyložiť a aplikovať prístupy, základné i moderné, ktoré sa využívajú pri odhade očakávanej hodnoty náhodnej veličiny X_{ij}. V matematickom ponímaní pôjde v tarifnej triede (i, j) o odhad $P_{i,j}$ definovaný ako

$$P_{ij} = EX_{ij}. \tag{1.1}$$

V prípade, že uvažujeme interpretáciu X_{ij} z bodu 3, tak odhad P_{ij} sa nazýva *nettopoistné*. Nettopoistné je čiastka, ktorá má byť zinkasovaná od jednotlivých poistencov za účelom pokrytia celkového poistného plnenia. Ak k nettopoistnému pripočítame položky, ktoré slúžia na pokrytie správnych nákladov a prípadné bezpečnostné prirážky, tak dostaneme *bruttopoistné*, ktoré sa niekedy nazýva i *poistným*. **V ďalšom texte budeme odhad P_{ij} nazývať už priamo poistným, nezávisle na tom o akú interpretáciu náhodných veličín X_{ij} pôjde.**

Budeme uvažovať dve základné modelové štruktúry, respektíve dva typy závislosti tarifných premenných

1. **ADITÍVNA ŠTRUKTÚRA**

$$P_{ij}^{A} = \mu_0 + \psi_i + \phi_j, \tag{1.2}$$

2. **MULTIPLIKATÍVNA ŠTRUKTÚRA**

$$P_{ij}^{M} = \mu_0 \psi_i \phi_j. \tag{1.3}$$

Poznámka 1.
μ_0 môžeme interpretovať napríklad ako

1. priemerné poistné a ψ_i, resp. ϕ_j predstavujú výchylky od priemerného poistného spôsobené vplyvom hodnoty tarifnej premennej A, resp. B.
2. poistné v dopredu zvolenej tarifnej triede, referenčnej triede. ψ_i, resp. ϕ_j reprezentujú výchylky od poistného v referenčnej triede spôsobené vplyvom hodnoty tarifnej premennej A, resp. B. Nech tarifná trieda (a, b) je referenčná trieda, potom pre prípad aditívnej štruktúry modelu je $\psi_a = 0$ a $\phi_b = 0$. V prípade multiplikatívnej štruktúry modelu je $\psi_a = 1$ a $\phi_b = 1$.

Úlohou nasledujúcich riadkov je odhadnúť hodnoty parametrov μ_0, ψ_i, ϕ_j na základe realizácií náhodných veličín X_{ij}, čím získame odhad P_{ij}, teda poistného pre $i = 1, 2, \ldots I, j = 1, 2, \ldots J$.

Obe štruktúry modelu možno zapísať všeobecne i pre viac tarifných premenných a to nasledovne.

1. **ADITÍVNA TARIFNÁ ŠTRUKTÚRA**

$$P_{k_1,k_2,\ldots k_K}^{A} = \mu_0 + \sum_{j=1}^{K} \alpha_{j,k_j}, \tag{1.4}$$

kde $\mu_0, \alpha_{j,k_j}, j = 1 \ldots K, k_j = 1 \ldots I_j$ sú reálne čísla.

2. **MULTIPLIKATÍVNA TARIFNÁ ŠTRUKTÚRA**

$$P_{k_1,k_2,\ldots k_K}^{M} = \mu_0 \prod_{j=1}^{K} \alpha_{j,k_j}, \tag{1.5}$$

kde $\mu_0, \alpha_{j,k_j}, j = 1 \ldots K, k_j = 1 \ldots I_j$ sú reálne čísla.

Kapitola 2

Základné metódy

Pod pojmom základné metódy budeme v tomto texte označovať:

1. **Vážená metóda najmenších štvorcov**
2. **Metóda marginálnych súčtov**
3. **Metóda Baileyho - Simonova**.

Postupne predvedieme tieto tri metódy aplikované na aditívnu i multiplikatívnu štruktúru.

2.1 Základné metódy a aditívna štruktúra

V celom tomto odstavci budeme pracovať s aditívnou tarifnou štruktúrou, ktorú opíšeme nasledovne.

Predpoklad 1.
Nech X_{ij}, $i = 1, 2, \dots I$ a $j = 1, 2, \dots J$ z 1.1 sú náhodné veličiny , ktoré spĺňajú vzťah

$$X_{ij} = \mu_0 + \psi_i + \phi_j + \epsilon_{ij}, \tag{2.1}$$

kde ϵ_{ij} sú nezávislé náhodné veličiny s

$$\begin{aligned} E\epsilon_{ij} &= 0, \\ var\epsilon_{ij} &= \frac{\sigma^2}{n_{ij}} \end{aligned} \tag{2.2}$$

a μ_0, ψ_i, ϕ_j sú reálne čísla.

Našou úlohou je na základe realizácií náhodných veličín X_{ij}, $i = 1, 2, \ldots I$, $j = 1, 2, \ldots J$ odhadnúť P_{ij}^A zavedené vzťahom 1.2.
Na to, aby sme vyčíslili poistné je nutné nájsť odhady parametrov μ_0, ψ_i, ϕ_j, ktoré si označíme $\hat{\mu_0}, \hat{\psi}_i, \hat{\phi}_j$, a tým dostaneme hľadané poistné P_{ij}^A v tvare

$$\hat{P_{ij}^A} = \hat{\mu_0} + \hat{\psi}_i + \hat{\phi}_j. \tag{2.3}$$

Poznámka 2.
V prípade interpretácie μ_0 ako priemerného poistného spĺňajú $\psi_i, \phi_j, i = 1, \ldots I, j = 1, \ldots J$ rovnosť

$$\sum_{i=1}^{I} n_{i\bullet}\psi_i + \sum_{j=1}^{J} n_{\bullet j}\phi_j = 0. \tag{2.4}$$

2.1.1 Vážená metóda najmenších štvorcov

Vážená metóda najmenších štvorcov (VMNŠ) je veľmi známa metóda, ktorá je založená na minimalizovaní váženého súčtu štvorcov rozdielu. V našom prípade rozdielu náhodných veličín X_{ij} a P_{ij}^A. Váhy budú v tomto prípade rovné miere expozície n_{ij}, teda napríklad počtu poistných zmluv v danej tarifnej triede.
Ide o minimalizáciu výrazu

$$Q = \sum_{i,j} n_{ij} \left(X_{ij} - P_{ij}^A \right)^2. \tag{2.5}$$

2.1.2 Metóda marginálnych súčtov

Základná myšlienka tejto metódy spočíva v tom, že pre veľké skupiny poistených by poistné malo byť rovné pozorovaným stratám. To možno vyjadriť rovnicami

$$\begin{aligned} \sum_{j=1}^{J} n_{ij}\hat{P_{ij}^A} &= \sum_{j=1}^{J} n_{ij}X_{ij}, \qquad i = 1, 2, \ldots I, \\ \sum_{i=1}^{I} n_{ij}\hat{P_{ij}^A} &= \sum_{i=1}^{I} n_{ij}X_{ij}, \qquad j = 1, 2, \ldots J. \end{aligned} \tag{2.6}$$

Veta 2.1.1.
Za Predpokladu 1 a pri interpretácií μ_0 ako priemerného poistného, dáva VMNŠ opísaná v predchádzajúcom odstavci a metóda marginálnych súčtov opísaná rov-

nicami 2.6 rovnaké výsledky, a to

$$\begin{aligned}
\hat{\mu}_0 &= \bar{X_{\bullet\bullet}}, \\
\hat{\psi}_i &= \left(\bar{X_{i\bullet}} - \bar{X_{\bullet\bullet}}\right) - \sum_{j=1}^{J} \frac{n_{ij}}{n_{i\bullet}} \hat{\phi}_j, \qquad i = 1, 2, \dots I, \\
\hat{\phi}_j &= \left(\bar{X_{\bullet j}} - \bar{X_{\bullet\bullet}}\right) - \sum_{i=1}^{I} \frac{n_{ij}}{n_{\bullet j}} \hat{\psi}_i, \qquad j = 1, 2, \dots J,
\end{aligned} \tag{2.7}$$

kde

$$\bar{X_{i\bullet}} = \sum_{j=1}^{J} \frac{n_{ij}}{n_{i\bullet}} X_{ij}, \qquad \bar{X_{\bullet j}} = \sum_{i=1}^{I} \frac{n_{ij}}{n_{\bullet j}} X_{ij}, \qquad \bar{X_{\bullet\bullet}} = \sum_{i,j} \frac{n_{ij}}{n_{\bullet\bullet}} X_{ij},$$

$$n_{i\bullet} = \sum_{j=1}^{J} n_{ij}, \qquad n_{\bullet j} = \sum_{i=1}^{I} n_{ij}, \qquad n_{\bullet\bullet} = \sum_{i,j} n_{ij}.$$

2.1.3 Metóda Baileyho - Simonova

Je metóda založená na minimalizácií výrazu

$$Q^* = \sum_{i,j} \frac{n_{ij}}{P_{ij}^A} \left(X_{ij} - P_{ij}^A\right)^2. \tag{2.8}$$

Tradičným postupom hľadania minima funkcie viacerých premenných dostaneme nasledujúce riešenie.

Veta 2.1.2.
Za Predpokladu 1, dáva metóda Baileyho Simona nasledujúce odhady parametrov.

$$\begin{aligned}
\sum_{i,j} \frac{n_{ij} X_{ij}^2}{\left(\hat{\mu}_0 + \hat{\psi}_i + \hat{\phi}_j\right)^2} &= \sum_{i,j} n_{ij}, \\
\sum_{j=1}^{J} \frac{n_{ij} X_{ij}^2}{\left(\hat{\mu}_0 + \hat{\psi}_i + \hat{\phi}_j\right)^2} &= \sum_{j=1}^{J} n_{ij}, \qquad i = 1, 2, \dots I, \\
\sum_{i=1}^{I} \frac{n_{ij} X_{ij}^2}{\left(\hat{\mu}_0 + \hat{\psi}_i + \hat{\phi}_j\right)^2} &= \sum_{i=1}^{I} n_{ij}, \qquad j = 1, 2, \dots J.
\end{aligned} \tag{2.9}$$

Poznámka 3.
Rovnice vo vetách 2.1.1 a 2.1.2 sa obyčajne riešia použitím iteračných metód.

2.2 Základné metódy a multiplikatívna štruktúra

Podobne ako sme rozobrali základné modely aplikované na aditívnu štruktúru, tak ich budeme aplikovať i na štruktúru multiplikatívnu, ktorú opíšeme nasledovne.

Predpoklad 2.
Nech náhodné veličiny X_{ij}, $i = 1 \dots I$ a $j = 1 \dots J$ z 1.1 splňujú vzťah

$$EX_{ij} = \mu_0 \psi_i \phi_j, \tag{2.10}$$

kde μ_0, ψ_i a ϕ_j sú reálne čísla.

Našou úlohou je opäť na základe realizácií X_{ij}, $i = 1 \dots I$ a $j = 1 \dots J$ odhadnúť μ_0, ψ_i a ϕ_j a tým získať odhad P_{ij}^M, $i = 1 \dots I$ a $j = 1 \dots J$ v tvare

$$\hat{P_{ij}^M} = \hat{\mu_0} \hat{\psi_i} \hat{\phi_j}. \tag{2.11}$$

Poznámka 4.
V prípade interpretácie μ_0 ako priemerného poistného musia ψ_i a ϕ_j, kde $i = 1, 2 \dots I$, $j = 1, 2 \dots J$ splňať rovnosť

$$\sum_{i,j} \frac{n_{ij} \psi_i \phi_j}{n_{\bullet\bullet}} = 1. \tag{2.12}$$

2.2.1 Vážená metóda najmenších štvorcov

Je založená na minimalizácií výrazu

$$Q = \sum_{i,j} n_{ij} \left(X_{ij} - P_{ij}^M \right)^2. \tag{2.13}$$

Veta 2.2.1.
Za Predpokladu 2 dostaneme VMNŠ nasledujúce odhady parametrov μ_0, ψ_i, ϕ_j, $i = 1, 2 \dots I$, $j = 1, 2 \dots J$.

$$\begin{aligned} \hat{\mu_0} &= \frac{\sum_{i,j} n_{ij} X_{ij} \hat{\psi_i} \hat{\phi_j}}{\sum_{i,j} n_{ij} \hat{\psi_i}^2 \hat{\phi_j}^2}, \\ \hat{\psi_i} &= \frac{1}{\hat{\mu_0}} \frac{\sum_{j=1}^J n_{ij} X_{ij} \hat{\phi_j}}{\sum_{j=1}^J n_{ij} \hat{\phi_j}^2}, \qquad i = 1, 2, \dots I, \\ \hat{\phi_j} &= \frac{1}{\hat{\mu_0}} \frac{\sum_{i=1}^I n_{ij} X_{ij} \hat{\psi_i}}{\sum_{i=1}^I n_{ij} \hat{\psi_i}^2}, \qquad j = 1, 2, \dots J. \end{aligned} \tag{2.14}$$

2.2.2 Metóda marginálnych súčtov

Na rozdiel od predchádzajúceho odstavca metóda marginálnych súčtov a VMNŠ aplikovaná na multiplikatívnu štruktúru nedajú rovnaké výsledky.
V prípade multiplikatívnej tarifnej štruktúry je základná myšlienka je taká istá ako v aditívnom prípade.

$$\begin{aligned}\sum_{j=1}^{J} n_{ij}\hat{P_{ij}^M} &= \sum_{j=1}^{J} n_{ij}X_{ij}, \qquad i=1,2,\dots I,\\ \sum_{i=1}^{I} n_{ij}\hat{P_{ij}^M} &= \sum_{i=1}^{I} n_{ij}X_{ij}, \qquad j=1,2,\dots J.\end{aligned} \tag{2.15}$$

Veta 2.2.2.
Za Predpokladu 2 a pri interpretácií μ_0 ako priemerného poistného dáva metóda marginálnych súčtov nasledujúce výsledky.

$$\begin{aligned}\hat{\mu_0} &= \bar{X_{\bullet\bullet}},\\ \hat{\psi_i}\left(\sum_{j=1}^{J} n_{ij}\hat{\phi_j}\right) &= \sum_{j=1}^{J} n_{ij}\frac{X_{ij}}{\bar{X_{\bullet\bullet}}}, \qquad i=1,\dots I,\\ \hat{\phi_j}\left(\sum_{i=1}^{I} n_{ij}\hat{\psi_i}\right) &= \sum_{i=1}^{I} n_{ij}\frac{X_{ij}}{\bar{X_{\bullet\bullet}}}, \qquad j=1,\dots J.\end{aligned} \tag{2.16}$$

2.2.3 Metóda Baileyho - Simonova

Je založená na minimalizácií výrazu

$$Q^* = \sum_{i,j} \frac{n_{ij}}{P_{ij}^M}\left(X_{ij} - P_{ij}^M\right)^2. \tag{2.17}$$

Veta 2.2.3.
Za Predpokladu 2 dáva metóda Baileyho - Simonova nasledujúce odhady parametrov.

$$\begin{aligned}\hat{\mu_0} &= \left(\sum_{i,j}\frac{n_{ij}X_{ij}^2}{\hat{\psi_i}\hat{\phi_j}} \Big/ \sum_{i,j} n_{ij}\hat{\psi_i}\hat{\phi_j}\right)^{1/2},\\ \hat{\psi_i} &= \left(\sum_{j=1}^{J}\frac{n_{ij}X_{ij}^2}{\hat{\mu_0}\hat{\phi_j}} \Big/ \sum_{j=1}^{J} n_{ij}\hat{\mu_0}\hat{\phi_j}\right)^{1/2}, \qquad i=1,\dots I,\\ \hat{\phi_j} &= \left(\sum_{i=1}^{I}\frac{n_{ij}X_{ij}^2}{\hat{\mu_0}\hat{\psi_i}} \Big/ \sum_{i=1}^{I} n_{ij}\hat{\mu_0}\hat{\psi_i}\right)^{1/2}, \qquad j=1,\dots J.\end{aligned} \tag{2.18}$$

Poznámka 5. Sústavy 2.14, 2.16, 2.18 je možné riešiť použitím iteračných metód.

Kapitola 3

Zobecnené lineárne modely

3.1 Stavba zobecnených lineárnych modelov

Zobecnené lineárne modely (Generalized Linear Models, GLM) je skupina modelov, ktoré zahrňujú známejšie tradičné linárne modely (TLM), analýzu rozptylu, logistické modely, loglineárne modely, multinomické modely a ďaľšie.

Tradičný lineárny model možno vyjadriť v tvare

$$\mathbf{X} = \mathbf{Y}\,\beta + \epsilon, \tag{3.1}$$

kde

- $\mathbf{X}$ je $D \times 1$ vektor pozorovaných náhodných veličín (vysvetľovaných premenných).
- $\mathbf{Y}$ je $D \times p$ matica vysvetľujúcich premenných, ktorej riadky zodpovedajú jednotlivým meraniam.
- β je $p \times 1$ vektor neznámych parametrov, ktoré vyjadrujú závislosť vysvetľovaných a vysvetľujúcich premenných a našou úlohou je ich odhadnúť.
- ϵ je $D \times 1$ vektor nezávislých normálne rozdelených náhodných veličín s $N(0, \sigma^2)$.

V týchto modeloch predpokladáme, že realizácia náhodného vektora $\mathbf{X}$ je kombinácia systematickej (nenáhodnej) zložky, reprezentovanej členom $\mathbf{Y}\beta$, a náhodnej zložky, reprezentovanej vektorom ϵ . Základom TLM je predpoklad, že ϵ je vektor nezávislých náhodných veličín s $N(0, \sigma^2)$. Základné výhody GLM oproti TLM sú založené na rozšírení predpokladu o rozdelení náhodného vektora $\mathbf{X}$ a o rozšírení predpokladu lineárneho vzťahu medzi pozorovanou náhodnou veličinou a parametrami modelu. Pomocou týchto rozšírení sa možno prostredníctvom modelov viac priblížiť k reálnym situáciam.

Prvým rozšírením TLM je predpoklad rozdelenia náhodných veličín $X_1, X_2, \dots X_D$. Predpokladáme, že pozorovaná náhodná veličina je veličina s rozdelením z exponenciálnej rodiny rozdelení.

Definícia 1.
Náhodná veličina **X** má rozdelenie **z rodiny exponenciálnych rozdelení** ak jej hustotu možno zapísať v tvare

$$f(x) = exp\left(\frac{x\,\theta - b\,(\theta)}{a\,(\phi)} + c\,(x, \phi)\right), \tag{3.2}$$

kde θ sa nazýva *kanonický parameter*, ϕ je *disperzný parameter*, $a\,(\phi)$ je spojitá kladná funkcia, $b\,(\theta)$ je dvakrát diferencovateľná konvexná funkcia a $c\,(x, \phi)$ je funkcia normujúca f, nezávislá na θ.

Do exponenciálnej rodiny rozdelení patria Normálne rozdelenie, Poissonovo rozdelenie, Binomické rozdelenie, Gamma rozdelenie, Inverzné Gaussovo rozdelenie a iné.
Prehľadnú tabuľku i s uvedením jednotlivých parametrov z Definície 1 nájdete v prílohe.

Ďalší rozdiel oproti tradičným lineárnym modelom je v tom, že v GLM vystupujú navyše

1. **Lineárny prediktor**, označíme η, je lineárna funkcia neznámych parametrov modelu.
2. **Spojovacia funkcia** (Link function), označíme g, prostá a diferencovateľná funkcia spájajúca systematickú zložku $\mathbf{Y}\beta$, označíme μ, s lineárnym prediktorom, η. V TLM platí $\mu = \eta$, teda spojovacia funckia g je identita.
3. Funcia zhŕňajúca ostatné vplyvy (offset function), označíme ξ.

Veľmi dôležitý je vzťah medzi lineárnym prediktorom $\eta = (\eta_1, \eta_2 \dots \eta_D)$ a spojovacou funkciou g, definovaný ako

$$g\,(\mu_i) = \eta_i \quad \Rightarrow \quad \mu_i = g^{-1}\,(\eta_i)\,, \quad i = 1, 2 \dots D \tag{3.3}$$

a

$$\eta_i = \sum_j y_{ij}\beta_j + \xi_i, \quad i = 1, 2 \dots D, \tag{3.4}$$

čo sa dá vyjadriť v maticovom tvare

$$\eta = \mathbf{Y}\beta + \xi. \tag{3.5}$$

Ak zvolíme spojovaciu funkciu nasledovne

$$\eta = g\left(\mu\right) = \theta, \tag{3.6}$$

bude lineárny prediktor priamo rovný kanonickému parametru z Definície 1. Takúto spojovaciu funkciu nazývame **kanonická spojovacia funkcia**.

V prípade GLM sa pri hľadaní odhadu neznámych parametrov $\beta = (\beta_1, \beta_2 \ldots, \beta_p)$ používa metóda maximálnej vierohodnosti, čo je jedna z klasických metód bodových odhadov. Nech náhodné veličiny $X_1, X_2, \ldots X_D$ majú rozdelenie charakterizované hustotami $f_i, i = 1 \ldots D$, ktoré poznáme až na vektor neznámych parametrov $\beta_1, \beta_2 \ldots, \beta_p$. Základná myšlienka metódy maximálnej vierohodnosti spočíva vo voľbe takých odhadoch neznámych parametrov, aby bola pravdepodobnosť realizácií náhodných veličín $X_1, X_2, \ldots X_D$ maximálna.

Prakticky metóda maximálnej vierohodnosti spočíva v hľadaní argumentu maxima **vierohodnostnej funkcie**, ktorú si označíme L, definovanej ako

$$L = \prod_{i=1}^{D} f\left(X_i, \beta_1, \beta_2, \ldots, \beta_p\right). \tag{3.7}$$

Často sa z technického dôvodu maximalizuje **logaritmická vierohodnostná funkcia**, definovaná ako

$$ln\ L = \sum_{i=1}^{D} ln\ f\left(X_i, \beta_1, \beta_2 \ldots \beta_p\right). \tag{3.8}$$

Poznámka 6.
Práve pomocou logaritmickej vierohodnostnej funkcie vieme odvodiť nasledujúce vzťahy.

$$EX = \acute{b}\left(\theta\right), \qquad varX = b''\left(\theta\right) a\left(\phi\right) = b''\left(\theta\right) \frac{\phi}{w}. \tag{3.9}$$

Zpravidla sa uvažuje, že $a\left(\phi\right) = \phi/w$, kde w je apriórna váha príslušného pozorovania.
K overeniu 3.9 stačí dosadiť do nasledujúcich známych vzťahov, ktoré platia pre náhodné veličiny s rozdelením z exponenciálnej rodiny rozdelení.

$$\begin{aligned} E\left(\frac{\partial ln\ L}{\partial \theta}\right) &= 0, \\ E\left(\frac{\partial^2 ln\ L}{\partial \theta^2}\right) + E\left(\frac{\partial ln\ L}{\partial \theta}\right)^2 &= 0. \end{aligned}$$

3.2 GLM v tarifovaní

Našou úlohou stále zostáva odhadnúť poistné pre jednotlivé tarifné triedy. Pripomenieme si označenie zavedené v predchádzajúcej kapitole a zavedieme to, s ktorým budeme pracovať v tejto kapitole.

Pripomenutie
V Kapitole 1 sme zaviedli nasledujúce označenie.

- X_{ij} je pozorovaná náhodná veličina v tarifnej triede (i, j).

- $P_{ij} = EX_{ij}$ je očakávaná hodnota pozorovanej náhodnej veličiny v tarifnej triede (i, j), teda poistné v tarifnej triede (i, j).

- μ_0 je poistné v referenčnej triede (alebo priemerné poistné).

- ψ_i je výchylka od poistného v referenčnej triede (alebo od priemerného poistného) μ_0, spôsobená vplyvom tarifnej premennej A.

- ϕ_j je výchylka od poistného v referenčnej triede (alebo od priemerného poistného) μ_0, spôsobená vplyvom tarifnej premennej B.

V 3.1 uvažujeme

- X_i, $i = 1, \ldots, D$, $D = I \times J$ je vektor zložený z náhodných veličín X_{ij}, $i = 1 \ldots I$, $j = 1 \ldots J$.

- $\beta_i, i = 1 \ldots p, p = I+J+1$ je vektor neznámych parametrov, ktorý vyjadruje vplyv zaradenia rizika do príslušnej tarifnej triedy na poistné (vzhľadom k poistnému v referenčnej triedy alebo priemernému poistnému) podobne ako predtým μ_0, ψ_i, $i = 1, 2 \ldots I$, ϕ_j, $j = 1, 2, \ldots J$.

Našou úlohou bude opäť na základe pozorovaní odhadnúť očakávanú (strednú) hodnotu pozorovaných náhodných veličín X_i, $i = 1, 2 \ldots D$.

$$P_i = EX_i, \qquad i = 1, 2 \ldots D. \tag{3.10}$$

Dôvody, pre ktoré nie sú pri odhade poistného vhodné TLM sú prioritne nasledovné.

- Predpoklad normality
  Nemá význam uvažovať, že počet škôd alebo veľkosť škody bude realizácia náhodnej veličiny s normálnym rozdelením. Nemožno totiž predpokladať, že počet škôd alebo veľkosť škody budú záporné. V GLM predpoklad normality nie je.

- Nelineárna štruktúra
  Lineárna štruktúra nám nezaručuje nezápornosť, ktorá je pre počet škôd a veľkosť škody istá. V GLM máme nástroj, spojovaciu funkciu.

- Nekonštantný rozptyl
  V tradičných lineárnych modeloch uvažujeme ten istý rozptyl pre každú tarifnú triedu, čo je veľmi silný predpoklad. Napríklad v prípade Poissonovho rozdelenia je stredná hodnota rovná rozptylu. V každej triede je však stredná hodnota iná a teda i rozptyl sa mení. V GLM tento predpoklad nemáme.

V nasledujúcich riadkoch si predvedieme aplikáciu GLM v tarifovaní. Konkrétne si ukážeme GLM v súvislosti s modelovaním počtu a výšky škôd.
Základom aplikácie GLM je vhodná voľba rozdelenia pozorovanej náhodnej veličiny.
V prípade modelovania počtu škôd sú najpoužívanejšie nasledovné.

1. **Poissonovo rozdelenie**: $Po(\lambda)$, $\lambda > 0$.
   Základné charakteristiky $Po(\lambda)$:

$$\begin{aligned} P(X=n) &= \frac{\lambda^n}{n!}e^{-\lambda}, \qquad n = 1, 2, \ldots \\ EX &= \lambda, \\ varX &= \lambda. \end{aligned} \tag{3.11}$$

2. **Negatívne binomické rozdelenie**: $NBi(h,p)$, $h > 0$, $p \in (0,1)$.
   Základné charakteristiky $NBi(h,p)$:

$$\begin{aligned} P(N=n) &= \binom{n+h-1}{n} p^h (1-p)^n, \qquad n = 0, 1, 2, \ldots \\ EX &= \frac{h(1-p)}{p}, \\ varX &= \frac{h(1-p)}{p^2}. \end{aligned} \tag{3.12}$$

V prípade modelovania veľkosti škôd prichádzajú do úvahy nasledujúce rozdelenia.

1. **Exponenciálne rozdelenie**: $Exp\,(\alpha)$, $\alpha > 0$.
   Toto rozdelenie nie je vhodné za predpokladu, že môžu nastať vysoké škody. V tomto prípade konverguje distribučná funkcia veľmi rýchlo k 0, čím môžeme podceniť možnosť, že nastane vysoká škoda.
   Základné charakteristiky $Exp\,(\alpha)$:
$$\begin{aligned} F\,(x) &= 1 - e^{-x/\alpha}, \qquad x > 0, \\ f\,(x) &= \frac{1}{\alpha} e^{-x/\alpha}, \qquad x > 0, \\ EX &= \alpha, \\ varX &= \alpha^2. \end{aligned} \tag{3.14}$$

2. **Paretovo rozdelenie**: $Pa\,(a, b)$, $a > 0$, $b > 0$.
   Na rozdiel od exponenciálneho rozdelenia sa Paretovo rozdelenie používa i k popisu vysokých škôd.
   Základné charakteristiky $Pa\,(a, b)$:
$$\begin{aligned} F\,(x) &= 1 - \left(\frac{b}{x}\right)^a, \qquad x \geq b, \\ f\,(x) &= \frac{a}{b}\left(\frac{b}{x}\right)^{a+1}, \qquad x \geq b, \\ EX &= \frac{a\,b}{a-1}, \qquad a > 1, \\ varX &= \frac{a\,b^2}{(a-1)^2\,(a-2)}, \qquad a > 2. \end{aligned} \tag{3.15}$$

   Problém neexistencie momentov možno odstrániť v prípade, že veľkosť škody omedzíme zhora konečnou hranicou. Výsledkom tejto úpravy je tzv. *Cenzorované Paretovo rozdelenie.*

3. **Gamma rozdelenie**: $\Gamma\,(a, p)$, $a > 0$, $p > 0$.

   Základné charakteristiky $\Gamma\,(a, p)$:
$$\begin{aligned} f\,(x) &= \frac{a^p}{\Gamma\,(p)} e^{-a\,x} x^{p-1}, \qquad x > 0, \\ EX &= \frac{p}{a}, \\ varX &= \frac{p}{a^2}. \end{aligned} \tag{3.16}$$

4. **Weibullovo rozdelenie**

5. **Logaritmicko normálne rozdelenie**

6. a iné

Dôvody pre voľbu práve týchto rozdelení možno nájsť napríklad v [5].

V tomto texte si vyberieme pre modelovanie počtu škôd Poissonovo rozdelenie a pre modelovanie veľkosti škôd Gamma rozdelenie. Obe tieto rozdelenia patria do rodiny exponenciálnych rozdelení.

Ďalšia voľba, s ktorou sa treba vysporiadať je voľba spojovacej funkcie. Jedna z možností je kanonická spojovacia funkcia definovaná vzťahom 3.6, ktorej výhodou je jednoduchosť tvaru logaritmickej vierohodnostnej funkcie. Na druhej strane je však interpretovateľnosť a praktickosť. V posťovníctve ja dobre interpretovateľná a najčastejšie používaná **logaritmická spojovacia funkcia**.

3.2.1 Modelovanie počtu škôd

To, že Poissonovo rozdelenie je skutočne rozdelenie z exponenciálnej rodiny rozdelení je zrejmé z nasledujúceho.
Nech X má $Po(\lambda)$, potom

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{(k \ln \lambda - \lambda)}}{k!}, \tag{3.17}$$

teda

$$\begin{aligned} \theta &= \ln \lambda, \\ b(\theta) &= e^{\theta}, \\ \phi &= 1, \\ w &= 1, \\ EX = \mu = \acute{b}(\theta) &= e^{\theta} = \lambda, \\ varX = \frac{b''(\theta)\,\phi}{w} &= e^{\theta} = \lambda. \end{aligned} \tag{3.18}$$

Pristúpme priamo k modelovaniu počtu škôd pomocou GLM.
Nech X_i je náhodná veličina, ktorá vyjadruje počet škôd v i-tej tarifnej triede a má Poissonovo rozdelenie s parametrom $\lambda_i n_i$.
Z predchádzajúceho vyplýva nasledujúce.

$$\begin{aligned} \theta_i &= \ln \lambda_i n_i, \\ EX_i = \mu_i &= e^{\theta_i} = \lambda_i n_i, \\ varX_i = e^{\theta_i} &= \lambda_i n_i. \end{aligned} \tag{3.19}$$

Volíme kanonickú spojovaciu funkciu, ktorá je rovná logaritmickej spojovacej funkcií.

$$\eta_i = \theta_i = ln\,\lambda_i n_i = ln\,\lambda_i + ln\,n_i = \sum_{j=1}^{p} y_{ij}\beta_j + \underbrace{ln\,n_i}_{\text{offset function}}. \tag{3.20}$$

Našou úlohou je odhadnúť parametre $\beta_1, \beta_2 \dots \beta_p$.
Vierohodnostná funkcia má tvar

$$L = \prod_{i=1}^{D} \frac{e^{\left(x_i\theta_i - e^{\theta_i}\right)}}{x_i!}. \tag{3.21}$$

Logaritmická vierohodnostná funkcia má tvar

$$ln\,L = \sum_{i=1}^{D} \left(x_i\theta_i - e^{\theta_i} - ln\,x_i!\right). \tag{3.22}$$

Dosadením 3.20 dostávame

$$ln\,L = \sum_{i=1}^{D} x_i \left(\sum_{j=1}^{p} y_{ij}\beta_j\right) + \sum_{i=1}^{D} x_i ln\,n_i - \sum_{i=1}^{D} n_i e^{\sum_{j=1}^{p} y_{ij}\beta_j} - \sum_{i=1}^{D} ln\,(x_i!). \tag{3.23}$$

Zderivujeme podľa jednotlivých neznámych parametrov, položíme rovno 0 a dostaneme p rovníc o p neznámych.
Ak odhady parametrov $\beta_1, \dots, \beta_p$ označíme $\hat{\beta}_1, \dots, \hat{\beta}_p$, tak možno písať

$$\hat{\theta}_i = \sum_{j=1}^{p} y_{ij}\hat{\beta}_j + ln\,n_i. \tag{3.24}$$

Poistné v jednotlivých tarifných triedach bude

$$P_i^{\hat{GLM}} = E\hat{X}_i = \hat{\lambda}_i n_i = g^{-1}\left(\hat{\theta}_i\right) = n_i e^{\sum_{i=1}^{p} y_{ij}\hat{\beta}_j}. \tag{3.25}$$

Tu je na mieste si uvedomiť, že poistné pre jednu poistnú zmluvu v príslušnej tarifnej triede je opäť v tvare súčinu troch prvkov, podobne ako tomu bolo pre prípad multiplikatívnej tarifnej štruktúry. Vzťah medzi parametrami, ktoré sa vyskytujú v 1.3 a parametrami v GLM možno opísať nasledujúcou tabuľkou.

Tabuľka 3.1: Označenie

Označenie z Kapitoly 2	**Označenie z Kapitoly 3**
I+J+1	p
μ_0	e^{β_1}
ψ_i, $i = 1, 2 \ldots I$	$e^{\beta_{i+1}}$
ϕ_j, $j = 1, 2 \ldots J$	$e^{\beta_{I+j+1}}$

3.2.2 Modelovanie veľkosti škôd

Pre prípad modelovania veľkosti škôd budeme uvažovať náhodnú veličinu X s Gamma rozdelením.Funkciu hustoty náhodnej veličiny X si označíme f.

Na tomto mieste je vhodné si upraviť parametre Gamma rozdelenia zavedené v 3.16 nasledovne.

$$\begin{aligned} \mu &= \frac{p}{a}, \\ \varphi &= p. \end{aligned} \tag{3.26}$$

Hustota náhodnej veličiny X s rozdelením $\Gamma(\mu, \varphi)$ má tvar

$$f(x) = \frac{1}{\Gamma(\varphi)} x^{\varphi - 1} e^{\frac{-\varphi\, x}{\mu}} \varphi^{\varphi} \mu^{-\varphi}. \tag{3.27}$$

To, že ide naozaj o rozdelenie z exponenciálnej rodiny rozdelení je zrejmé z nasledujúcej úpravy 3.27.

$$f(x) = exp\left(\frac{-\frac{x}{\mu} - ln\ \mu}{\frac{1}{\varphi}} + \varphi\, ln\ (\varphi x) - lnx - ln\, \Gamma(\varphi) \right). \tag{3.28}$$

Teda:

$$\begin{aligned}
\theta &= \frac{-1}{\mu}, \\
b(\theta) &= -ln\ (-\theta), \\
\phi &= \frac{1}{\varphi}, \\
a(\phi) &= \frac{1}{\varphi}, \\
w &= 1, \\
EX &= \frac{-1}{\theta} = \mu, \\
varX &= \frac{\phi}{\theta^2} = \frac{1}{\theta^2 \varphi} = \frac{\mu^2}{\varphi}.
\end{aligned} \tag{3.29}$$

Tvrdenie 3.2.1.
Nech $X_1, X_2, \dots X_n$ sú nezávislé náhodné veličiny s rozdelením $\Gamma(\mu, \varphi)$. Nech Y je náhodná veličina, definovaná ako $Y = \sum_{i=1}^{n} X_i$. Potom náhodná veličina Y má rozdelenie $\Gamma(n\mu, n\varphi)$.

Budeme predpokladať, že jednotlivé dieľčie škody v tarifnej triede i sú realizície náhodných veličín s $\Gamma(\mu_i, \varphi_i)$. Potom úhrn všetkých škôd,n_i, možno považovať za realizáciu náhodných veličín s rozdelením $\Gamma(n_i\mu_i, n_i\varphi_i)$.

V prípade modelovania veľkosti škôd má kanonická spojovacia funkcia tvar

$$\eta_i = \theta_i = -\frac{1}{n_i \mu_i} = -\sum_{j=1}^{p} y_{ij}\beta_j. \tag{3.30}$$

V tomto prípade však kanonická spojovacia funkcia má nevýhodu v tom, že negarantuje nezápornosť strednej hodnoty.
Ako bolo naznačené tak v poisťovníctve sa ukázalo byť užitočné použitie logaritmickej spojovacej funkcie. I keď technicky to neznamená zjednodušenie funkcie maximálnej vierohodnosti na druhej strane použitie logaritmickej spojovacej funkcie nevedie k negatívne odhadnutej strednej hodnote. K aplikácií GLM na modelovanie veľkosti škôd použijeme logaritmickú spojovaciu funkciu

$$\eta_i = ln\ (n_i\mu_i) = ln\left(-\frac{1}{\theta_i}\right) = ln\ n_i + ln\ \mu_i = \sum_{j=1}^{p} y_{ij}\beta_j + ln\ n_i. \tag{3.31}$$

Teda

$$\theta_i = -\frac{1}{n_i\mu_i} = -\frac{e^{\left(-\sum_{j=1}^{p} y_{ij}\beta_j\right)}}{n_i}. \tag{3.32}$$

Podobne ako v prípade modelovania počtu škôd možno použiť k odhadu neznámych parametrov $\beta_1, \beta_2 \dots \beta_p$ metódu maximálnej vierohodnosti.
Poistné zapíšeme podobne ako v 3.25.

3.3 Overenie platnosti modelu

Nástrojov na overenie platnosti modelov je niekoľko. I v prípade GLM je veľmi obľúbené testovanie hypotéz. To ako veľmi sa realizície náhodných veličín líšia od hodnôt odhadnutých možno skúmať pomocou **deviancie** a **Pearsonovej χ^2 štatistiky**.

1. DEVIANCIA $D(x, \mu)$
   Definujeme je pomocou logaritmickej vierohodnostnej funkcie nasledovne.
   $$D(x, \mu)/\phi = 2(ln\ L(x, x) - ln\ L(x, \mu))/\phi, \tag{3.33}$$
   kde $x = (x_1, x_2 \dots x_D)$ sú realizácie pozorovaných náhodných veličín, $\mu = (\mu_1, \mu_2, \dots \mu_D)$ sú modelom získané odhady stredných hodnôt náhodných veličín, L je vierohodnostná funkcia a ϕ je disperzný parameter.
   Pre pevné ϕ definujeme tzv. **skalárnu devianciu**
   $$D^*(x, \mu) = D(x, \mu)/\phi. \tag{3.34}$$

2. Pearsonova χ^2 štatistika
   Je daná vzťahom
   $$Q = \sum_{i=1}^{D} \frac{(x_i - \mu_i)^2}{V_{\mu_i}}, \tag{3.35}$$
   kde V_{μ_i} je odhad rozptylu v tarifnej triede i, $i = 1 \dots D$. Skalárna verzia Pearsonovej χ^2 štatistiky je definovaná ako
   $$Q^* = Q/\phi. \tag{3.36}$$

Obe tieto štatistiky, deviancia i Pearsonova χ^2 štatistika, majú χ^2 rozdelenie o počte stupňov voľnosti rovnému rozdielu medzi počtom pozorovaní a počtu odhadovaných parametrov.

Ďalšou možnosťou ako vyšetriť vhodnosť modelu je preskúmať ako sa chovajú reziduá. V TLM sa najčastejšie používajú reziduá $x - \mu$. V GLM tieto typické reziduá využívať nemožno, lebo $VarX_i \neq kon.$ pre každé i. Z tohto dôvodu boli zavedené iné reziduá ako napríklad **deviančné reziduum** a **Paersonove reziduum**.

1. Deviančné reziduá definujeme nasledovne.
   $$d_{i,r} = [sgn(x_i - \mu_i)]\sqrt{d_i}, \tag{3.37}$$
   kde d_i je definované vzťahom
   $$\sum_{i=1}^{D} d_i = D(x, \mu). \tag{3.38}$$

2. Pearsonove reziduá definujeme nasledovne

$$r_{i,p} = \frac{x_i - \mu_i}{\sqrt{V_{\mu_i}}}. \tag{3.39}$$

Kapitola 4

Kredibilitné modely

V posledných rokoch sa zvyklo zavádzať čoraz viac tarifných premenných s cieľom definovania rizika, čo najlepšie, ako je to možné. Nevýhoda stáleho pribúdania tarifných premenných je v tom, že čím viac ich zavedieme, tým menej údajov pre jednotlivé tarifné triedy dostaneme.
Metóda GLM nám poskytuje bodové odhady v jednotlivých tarifných triedach. Problém nastáva pokiaľ pri bodových odhadoch vychádzame z malého počtu pozorovaní. Vtedy už bodový odhad môže byť výrazne nepresný. Tu je namieste dať priestor nielen odhadom zisteným na základe dát v danej tarifnej triede, ale zahrnúť do poistného i vplyv celého poistného kmeňa. Práve kombinácia týchto dvoch zložiek, individuálnej a kolektívnej, je výsledkom odhadov, ktoré sa nazývajú **kredibilitné odhady**.

Definícia 2.
Povieme, že odhad náhodnej veličiny P, ktorý si označíme $\hat{P}$, je lepší ako iný odhad tej istej náhodnej veličiny P, označíme $\tilde{P}$, ak

$$E\left[\left(P-\hat{P}\right)^2\right] \leq E\left[\left(P-\tilde{P}\right)^2\right]. \tag{4.1}$$

Poznámka 7.
Najlepší odhad poistného $\hat{P_{ij}}$ v zmysle Definície 2 minimalizuje strednú kvadratickú chybu

$$E\left[\left(P-\hat{P}\right)^2\right]. \tag{4.2}$$

Kredibilitné odhady sú špeciálnym prípadom **Bayesovských odhadov**. Bayesovský odhad je najlepší odhad v zmysle Poznámky 7, kde uvažujeme, že parameter, ktorý odhadujeme je realizácia náhodnej veličiny s nejakým apriórnym rozdelením.
Nevýhodou Bayesovských odhadov je, že nie je väčšinou možné nájsť ich jednoduchý analytický tvar a je pri riešení úlohy nutné použiť rôzne matematické

postupy. Práve z tohoto dôvodu sa zavádza pojem **kredibilitný odhad**, čo je **najlepší lineárny odhad**, v zmysle Poznámky 7. Kredibilitné odhady sa z tohto dôvodu niekedy nazývajú i *lineárne bayesovské odhady* a budeme ich označovať horným indexom C (credibility).
Obecne sa kredibilitné odhady definujú nasledovne.

Definícia 3.
Kredilitný odhad parametru P je najlepší odhad, založený na všetkých pozorovaniach daných vektorom $X = (X_j, j = 1, 2, \ldots)$, ktorý je z priestoru

$$L(X, 1) := \left\{ \hat{P} : \hat{P} = a_0 + \sum_j a_j X_j; a_0, a_1 \ldots \in R \right\}. \tag{4.3}$$

Teda

$$P^C = \arg min_{\hat{P} \in L(\mathbf{X},1)} E\left[\left(\hat{P} - P \right)^2 \right].$$

Na rozdiel od vzťahov 1.2 a 1.3 je P^C náhodná veličina.

K hľadaniu kredibilitných odhadov sa často pristupuje elegantným spôsobom založeným na ortogonálnej projekcií neznámeho parametru do vhodného podpriestoru priestoru náhodných veličín s konečnými druhými momentami, L^2, čo je Hilbertov priestor. Priestor $L(X, 1)$ definovaný v Definícií 3 je podpriestorom v L^2. Pomocou tohoto prístupu možno kredibilitný odhad parametru P zapísať ako

$$P^C = Pro\left(P | L(\mathbf{X}, 1)\right). \tag{4.4}$$

Teraz si uvedieme dve tvrdenia, ktoré v ďalšom texte použijeme pri hľadaní kredibilitných odhadov.

Tvrdenie 4.0.1.
Nech M a $\acute{M}$ su uzavreté podpriestory Hilbertovho priestoru L^2 a $M \subset \acute{M}$. Potom platí

$$Pro(Y|M) = Pro\left(Pro\left(Y|\acute{M}\right)|M\right) \tag{4.5}$$

a

$$\|Y - Pro(Y|M)\|^2 = \left\|Y - Pro\left(Y|\acute{M}\right)\right\|^2 + \left\|Pro\left(Y|\acute{M}\right) - Pro(Y|M)\right\|^2, \tag{4.6}$$

kde $\|\cdot\|$ je norma.

Tvrdenie 4.0.2. ***Normálové rovnice***
P^C je kredibilitný odhad parametru P založený na všetkých pozorovaniach daných vektorom $X = (X_1, X_2, \dots X_n)$ práve vtedy, keď platia nasledujúce rovnice.

$$\begin{aligned} E\left(P^C - P\right) &= 0, \\ Cov\left(P, X_k\right) &= Cov\left(P^C, X_k\right), \qquad k = 1, 2, \dots n. \end{aligned} \tag{4.7}$$

Poznámka 8.
Pre P^C v tvare $P^C = \hat{a_0} + \sum_{j=1}^{n} \hat{a}_j X_j$ majú normálové rovnice definované vzťahmi 4.7 tvar

$$\begin{aligned} \hat{a_0} &= E\left(P\right) - \sum_j \hat{a}_j \mu_{X_j}, \\ \sum_j \hat{a}_j Cov\left(X_j, X_k\right) &= Cov\left(P, X_k\right), \qquad k = 1, 2, \dots n, \end{aligned} \tag{4.8}$$

kde $\mu_{X_j} = EX_j$.

Postupne v nasledujúcich dvoch odstavcov odvodíme kredibilitné odhady poistného pre prípad aditívnej i multiplikatívnej tarifnej štruktúry.

4.1 Kredibilitné odhady v aditívnej tarifnej štruktúre

Bayesovský aditívny model definujeme nasledovne.

Predpoklad 3.

1. Náhodné veličiny X_{ij}, $i = 1, \dots I$, $j = 1 \dots J$ sú za podmienky $\Theta_{ij} = (\Psi_i, \Phi_j)$ nezávislé náhodné veličiny s

$$\begin{aligned} P_{ij}^A\left(\Theta\right) &= E\left[X_{ij}|\Theta_{ij}\right] = \mu_0 + \Psi_i + \Phi_j, \\ Var\left(X_{ij}|\Theta_{ij}\right) &= \frac{\sigma^2\left(\Theta_{ij}\right)}{n_{ij}}. \end{aligned} \tag{4.9}$$

2. Náhodné veličiny $\Psi_i, i = 1, \dots I$ sú nezávislé, rovnako rozdelené náhodné veličiny s

$$\begin{aligned} E\left[\Psi_i\right] &= 0, \\ Var\left(\Psi_i\right) &= \tau_\Psi^2. \end{aligned} \tag{4.10}$$

3. Náhodné veličiny $\Phi_j, j = 1, \ldots J$ sú nezávislé, rovnako rozdelené náhodné veličiny s

$$\begin{aligned} E\left[\Phi_j\right] &= 0, \\ Var\left(\Phi_j\right) &= \tau_\Phi^2. \end{aligned} \tag{4.11}$$

4. $\Psi_i, \Phi_j, i = 1, \ldots I, j = 1 \ldots J$ sú nezávislé náhodné veličiny.

Kredibilitný odhad parametru v prípade aditívnej tarifnej štruktúry budeme označovať horným indexom CA.

Veta 4.1.1. *1. Za Predpokladu 3 má kredibilitný odhad P_{ij}^A, ktorý označíme $P_{ij}^{CA}(\Theta)$ tvar*

$$P_{ij}^{CA}(\Theta) = \mu_0 + \Psi_i^{CA} + \Phi_j^{CA}, \tag{4.12}$$

kde Ψ_i^{CA}, Φ_j^{CA} sú kredibilitné odhady veličín Ψ_i, Φ_j.

2. Ψ_i^{CA} a Φ_j^{CA} sú dané vzťahmi

$$\begin{aligned} \Psi_i^{CA} &= \alpha_i\left(\bar{X_{i\bullet}} - \mu_0\right) - \alpha_i \sum_{j=1}^{J} \frac{n_{ij}}{n_{i\bullet}} \Phi_j^{CA}, \end{aligned} \tag{4.13}$$

$$\begin{aligned} \Phi_j^{CA} &= \beta_j\left(\bar{X_{\bullet j}} - \mu_0\right) - \beta_j \sum_{i=1}^{I} \frac{n_{ij}}{n_{\bullet j}} \Psi_i^{CA}, \end{aligned} \tag{4.14}$$

kde

$$\begin{aligned} \alpha_i &= \frac{n_{i\bullet}}{n_{i\bullet} + \frac{\sigma^2}{\tau_\Psi^2}}, \\ \beta_j &= \frac{n_{\bullet j}}{n_{\bullet j} + \frac{\sigma^2}{\tau_\Phi^2}}, \\ \sigma^2 &= E\left[\sigma^2\left(\Theta_{ij}\right)\right]. \end{aligned} \tag{4.15}$$

Dôkaz

1. Rovnica 4.12 je zrejmá, keďže kredibilitný odhad sa vyznačuje linearitou. Ak nájdeme kredibilitné odhady parametrov Ψ_i, Φ_j dokážeme vyjadriť kredibilitný odhad P_{ij}^A.

2. Použitím 4.0.1 píšeme

$$\Psi_i^{CA} = Pro\left(\Psi_i | L\left(\mathbf{X}, 1\right)\right) = Pro\left(\underbrace{Pro\left(\Psi | L\left(X, 1, \mathbf{\Phi}\right)\right)}_{\Psi_i^*} | L\left(\mathbf{X}, 1\right)\right). \tag{4.16}$$

Ukážeme, že pre Ψ_i^* platí

$$\Psi_i^* = \alpha_i \left(\bar{X_{i\bullet}} - \mu_0 \right) - \alpha_i \sum_{j=1}^{J} \frac{n_{ij}}{n_{i\bullet}} \Phi_j. \tag{4.17}$$

Overíme, či sú splnené normálové rovnice z Tvrdenia 4.0.2. Teda overíme rovnosti

$$\begin{aligned} E\left(\Psi_i^* - \Psi_i\right) &= 0, && (4.18)\\ Cov\left(\Psi_i^*, X_{kl}\right) &= Cov\left(\Psi_i, X_{kl}\right), && (4.19)\\ Cov\left(\Psi_i^*, \Phi_l\right) &= Cov\left(\Psi_i, \Phi_l\right). && (4.20) \end{aligned}$$

S použitím Predpokladov 3 píšeme
Pre 4.18

$$\begin{aligned} E\left(\Psi_i^*\right) &= \alpha_i \underbrace{E\left(\left(\bar{X_{i\bullet}} - \mu_0\right)\right)}_{=0} - \alpha_i \sum_{j=1}^{J} \frac{n_{ij}}{n_{i\bullet}} \underbrace{E\left(\Phi_j\right)}_{=0} = 0,\\ E\left(\Psi_i\right) &= 0. \end{aligned}$$

Pre 4.19

$$\begin{aligned} Cov\left(\Psi_i^*, X_{kl}\right) &= \underbrace{\alpha_i \sum_j \frac{n_{ij}}{n_{i\bullet}} Cov\left(X_{ij}, X_{kl}\right)}_{=\frac{\alpha_i}{n_{i\bullet}}\sigma^2\delta_{ik}+\alpha_i\tau^2\Psi\delta_{ik}+\alpha_i\frac{n_{il}}{n_{i\bullet}}\tau^2_\Phi} + \alpha_i \sum_j \frac{n_{ij}}{n_{i\bullet}} \underbrace{Cov\left(\Phi_j, X_{kl}\right)}_{=\tau^2_\Phi\delta_{jl}} =\\ &= \frac{\alpha_i}{n_{i\bullet}}\sigma^2\delta_{ik} + \alpha_i\tau^2_\Psi\delta_{ik} + \alpha_i\frac{n_{il}}{n_{i\bullet}}\tau^2\Phi - \alpha_i\frac{n_{il}}{n_{i\bullet}}\tau^2_\Phi =\\ &= \alpha_i\left(\frac{\sigma^2}{n_{i\bullet} + \tau^2_\Psi}\right)\delta_{ik} = \tau^2_\Psi\delta_{ik},\\ Cov\left(\Psi_i, X_{kl}\right) &= E\Psi_i X_{kl} - \underbrace{E\Psi_i E X_{kl}}_{=0} = E\Psi_i\left(\mu_0 + \Psi_k + \Phi_l\right) =\\ &= \underbrace{\mu_0 E\Psi_i}_{=0} + E\Psi_i\Psi_k + \underbrace{E\Psi_i\Phi_l}_{=0} = \tau^2_\Psi\delta_{ik} \end{aligned}$$

Pre 4.20

$$\begin{aligned} Cov\left(\Psi_i^*, \Phi_l\right) &= \alpha_i \sum_j \frac{n_{ij}}{n_{i\bullet}} \underbrace{Cov\left(X_{ij}, \Phi_l\right)}_{=\tau^2_\Phi\delta_{jl}} - \alpha_i \sum_j \frac{n_{ij}}{n_{i\bullet}} \underbrace{Cov\left(\Phi_j, \Phi_l\right)}_{=\tau^2_\Phi\delta_{jl}} =\\ &= \alpha_i\frac{n_{il}}{n_{i\bullet}}\tau^2_\Phi - \alpha_i\frac{n_{il}}{n_{i\bullet}}\tau^2_\Phi = 0\\ &= Cov\left(\Psi_i, \Phi_l\right) = 0. \end{aligned}$$

Týmto sme overili vzťahy 4.18 - 4.20. Dosadením 4.17 do 4.16 dostávame

$$\begin{aligned}\Psi_i^{CA} &= \alpha_i\left(\bar{X}_{i\bullet}-\mu_0\right)-\alpha_i\sum_{j=1}^{J}\frac{n_{ij}}{n_{i\bullet}}Pro\left(\Phi_j|L\left(\mathbf{X},1\right)\right)= \\ &= \alpha_i\left(X_{i\bullet}-\mu_0\right)-\alpha_i\sum_{j=1}^{J}\frac{n_{ij}}{n_{i\bullet}}\Phi_j^{CA}.\end{aligned}$$

Podobne by sme dokázali i 4.14.

Na začiatku tejto kapitoly sme uviedli, že kredibilitné odhady poistného v tarifnej triede (i,j) na rozdiel od odhadov, ktoré získame metódou GLM, obsahujú i zložku, ktorá vyjadruje vplyv celého poistného kmeňa na poistné v triede (i,j), nie iba zmluv v tarifnej triede (i,j). Túto vlastnosť si predvedieme v nasledujúcom tvrdení, v ktorom uvažujeme pre jednoduchosť $n_{ij}=1$, $i=1\dots I$, $j=1,\dots J$.

Tvrdenie 4.1.2.
Nech platí Predpoklad 3 a nech $n_{ij}=1$, $i=1\dots I$, $j=1,\dots J$. Kredibilitné odhady poistného P_{ij}^{A} majú tvar

$$P_{ij}^{CA}\left(\Theta\right)=\mu_0+\alpha\left(\bar{X}_{i\bullet}-\mu_0\right)+\beta\left(\bar{X}_{\bullet j}-\mu_0\right)-\gamma\left(\bar{X_{\bullet\bullet}}-\mu_0\right), \tag{4.21}$$

kde

$$\alpha=\frac{J}{J+\frac{\sigma^2}{\tau_\Psi^2}},\qquad \beta=\frac{I}{I+\frac{\sigma^2}{\tau_\Phi^2}},\qquad \gamma=\frac{\alpha\beta}{1-\alpha\beta}\left(2-\alpha\beta\right).$$

V 4.21 na pravej strane máme 4 sčítance. Prvým sčítancom je priemerné poistné (alebo poistné v referenčnej tarifnej triede), druhý vyjadruje vplyv tarifnej premennej A, tretí vplyv tarifnej premennej B a štvrtý sčítanec vyjadruje korekciu založenú na dátach v celom poistnom kmeni.

<u>Dôkaz</u>
Vzťah 4.21 možno prepísať nasledovne:

$$P_{ij}^{CA}=\mu_0+\Psi_i^{CA}+\Phi_j^{CA}, \tag{4.22}$$

kde

$$\begin{aligned}\Psi_i^{CA} &= \alpha\left(\bar{X}_{i\bullet}-\mu_0\right)-\frac{\alpha\beta}{1-\alpha\beta}\left(1-\alpha\right)\left(\bar{X}_{\bullet\bullet}-\mu_0\right), \\ \Phi_j^{CA} &= \beta\left(\bar{X}_{\bullet j}-\mu_0\right)-\frac{\alpha\beta}{1-\alpha\beta}\left(1-\beta\right)\left(\bar{X}_{\bullet\bullet}-\mu_0\right).\end{aligned}$$

To súhlasí s 4.13 a 4.14. Stačí ak dosadíme:

$$\Psi_i^{CA} = \alpha\left(\bar{X_{i\bullet}} - \mu_0\right) - \alpha \sum_{j=1}^{J} \frac{1}{J} \Phi_i^{CA} =$$

$$\alpha\left(\bar{X_{i\bullet}} - \mu_0\right) - \alpha \sum_{j=1}^{J} \frac{1}{J} \left(\beta\left(\bar{X_{\bullet j}} - \mu_0\right) - \frac{\alpha\beta}{1-\alpha\beta}(1-\beta)\left(\bar{X_{\bullet\bullet}} - \mu_0\right)\right) =$$

$$\alpha\left(\bar{X_{i\bullet}} - \mu_0\right) - \alpha\beta \left(\underbrace{\sum_{j-1}^{J} \frac{1}{J} X_{\bullet j}}_{=\bar{X_{\bullet\bullet}}} - \mu_0\right) + \alpha\beta\left(\bar{X_{\bullet\bullet}} - \mu_0\right) \underbrace{\left[\frac{\alpha(1-\beta)}{1-\alpha\beta} - 1\right]}_{=\frac{\alpha-\alpha\beta-1+\alpha\beta}{1-\alpha\beta}} =$$

$$\alpha\left(\bar{X}_{i\bullet} - \mu_0\right) - \frac{\alpha\beta}{1-\alpha\beta}(1-\alpha)\left(\bar{X}_{\bullet\bullet} - \mu_0\right).$$

Podobne by sa dokazoval i vzťah pre Φ_j^{CA}.

4.2 Kredibilitné odhady v multiplikatívnej tarifnej štruktúre

Bayesovský multiplikatívny model definujeme nasledovne.

Predpoklad 4.

1. Náhodné veličiny X_{ij}, $i = 1, \ldots I$, $j = 1 \ldots J$ sú za podmienky $\Theta_{ij} = (\Psi_i, \Phi_j)$ nezávislé náhodné veličiny s

$$\begin{aligned} P_{ij}^{M}(\Theta) &= E\left[X_{ij}|\Theta_{ij}\right] = \mu_0 \Psi_i \Phi_j, \\ Var\left(X_{ij}|\Theta_{ij}\right) &= \frac{\sigma^2\left(\Theta_{ij}\right)}{n_{ij}} = \frac{\phi\left(\mu_0 \Psi_i \Phi_j\right)^p}{n_{ij}}, \end{aligned} \tag{4.23}$$

   kde ϕ a $p \in R^+$.

2. Náhodné veličiny $\Psi_i, i = 1, \ldots I$ sú nezávislé, rovnako rozdelené náhodné veličiny s

$$\begin{aligned} E\left[\Psi_i\right] &= 1, \\ Var\left(\Psi_i\right) &= \tau_\Psi^2. \end{aligned} \tag{4.24}$$

3. Náhodné veličiny $\Phi_j, j = 1, \ldots J$ sú nezávislé, rovnako rozdelené náhodné veličiny s

$$\begin{aligned} E\left[\Phi_j\right] &= 1, \\ Var\left(\Phi_j\right) &= \tau_\Phi^2. \end{aligned} \tag{4.25}$$

4. Ψ_i, Φ_j, $i = 1, \ldots I$, $j = 1 \ldots J$ sú nezávislé náhodné veličiny.

Poznámka 9.
ϕ z Predpokladu 4 sa nazýva disperzný parameter. Pre Poissonovo rozdelenie je $p = 1$ a pre Gamma rozdelenie je $p = 2$.

Z predchádzajúceho vieme, že kredibilitný odhad je najlepší lineárny odhad v zmysle strednej kvadratickej chyby. Lineárna štruktúra odhadu ovšem v tomto prípade nie je úplne najšťastnejšia. Preto musíme nájsť inú, vhodnejšiu stavbu odhadu.

Definícia 4.
Nech $\Psi_i^* (\mathbf{\Phi})$ respektíve $\Phi_j^* (\mathbf{\Psi})$, sú kredibilitné odhady Ψ_i, respektíve Φ_j, za podmienky $\mathbf{\Psi} = (\Psi_1, \Psi_2 \ldots \Psi_I)$, respektíve $\mathbf{\Phi} = (\Phi_1, \Phi_2 \ldots \Phi_J)$.

Parametre $\Psi_i^* (\mathbf{\Phi})$ a $\Phi_j^* (\mathbf{\Psi})$ sú závislé na parametroch Ψ_i, $i = 1 \ldots I$, Φ_j, $j = 1 \ldots J$. Takýto typ odhadov nazývame **pseudoodhady**.

Pre úplnosť si uvedieme predpoklady a závery Bühlmann Straubovho modelu, ktorý využijeme v nasledujúcom dôkaze.

Tvrdenie 4.2.1. ***Bühlmann Straubov model***
Nech máme riziká i, $i = 1 \ldots I$, ktoré sú charakterizované rizikovým parametrom θ_i, ktorý je realizáciou náhodnej veličiny Θ_i a nech súčastne platí nasledovné.

1. *Za podmienky Θ_i sú $\{X_{ij} : j = 1 \ldots J\}$ nezávislé náhodné veličiny s*

$$\begin{aligned} E\left[X_{ij}|\Theta_i, n_{ij}\right] &= \mu\left(\Theta_i\right), \\ Var\left[X_{ij}|\Theta_i, n_{ij}\right] &= \frac{\sigma^2\left(\Theta_i\right)}{n_{ij}}. \end{aligned}$$

2. *Dvojice $(\Theta_1, \mathbf{X_1}), (\Theta_2, \mathbf{X_2}) \ldots (\Theta_I, \mathbf{X_I})$ sú nezávislé a náhodné veličiny $\Theta_1, \Theta_2 \ldots \Theta_I$ sú nezávislé a rovnako rozdelené náhodné veličiny.*

Potom kredibilitný odhad $\mu(\Theta_i)$, $i = 1 \ldots I$ má tvar

$$\mu^C\left(\Theta_i\right) = \alpha_i \bar{X}_{i\bullet} + (1 - \alpha)\,\mu = \mu + \alpha_i\left(\bar{X}_{i\bullet} - \mu\right), \tag{4.26}$$

kde

$$\begin{aligned}
\bar{X}_{i\bullet} &= \sum_{j=1}^{J} \frac{n_{ij}}{n_{i\bullet}} X_{ij}, \\
n_{i\bullet} &= \sum_{j=1}^{J} n_{ij}, \\
\alpha_i &= \frac{n_{i\bullet}}{n_{i\bullet + \frac{\sigma^2}{\tau^2}}}, \\
\sigma^2 &= E\left[\sigma^2\left(\Theta_i\right)\right], \\
\tau^2 &= Var\left[\mu\left(\Theta_i\right)\right], \\
\mu &= E\left[\mu\left(\Theta_i\right)\right].
\end{aligned}$$

Veta 4.2.2.
Za Predpokladu 4, majú $\Psi_i^*(\mathbf{\Phi})$ a $\Phi_j^*(\mathbf{\Psi})$ tvar

$$\begin{aligned}
\Psi_i^*(\mathbf{\Phi}) &= 1 + \alpha_i \left(\bar{X}_{i\bullet}^{(1)} - 1\right), \\
\Phi_j^*(\mathbf{\Psi}) &= 1 + \beta_j \left(\bar{X}_{\bullet j}^{(2)} - 1\right),
\end{aligned} \tag{4.27}$$

kde

$$\begin{aligned}
X_{ij}^{(1)} &= \frac{X_{ij}}{\Phi_j \mu_0}, \\
n_{ij}^{(1)} &= n_{ij} \left(\Phi_j \mu_0\right)^{2-p}, \\
\bar{X}_{i\bullet}^{(1)} &= \sum_{j=1}^{J} \frac{n_{ij}^{(1)}}{n_{i\bullet}^{(1)}} X_{ij}^{(1)}, \\
\alpha_i &= \frac{n_{i\bullet}^{(1)}}{n_{i\bullet}^{(1)} + \frac{\sigma_\Psi^2}{\tau_\Psi^2}} \\
\sigma_\Psi^2 &= \eta \cdot E\left[\Psi_i^p\right], \\
X_{ij}^{(2)} &= \frac{X_{ij}}{\Psi_i \mu_0}, \\
n_{ij}^{(2)} &= n_{ij} \left(\Psi_i \mu_0\right)^{2-p}, \\
\bar{X}_{\bullet j}^{(2)} &= \sum_{i=1}^{I} \frac{n_{ij}^{(2)}}{n_{\bullet j}^{(2)}} X_{ij}^{(2)}, \\
\beta_j &= \frac{n_{\bullet j}^{(2)}}{n_{\bullet j}^{(1)} + \frac{\sigma_\Phi^2}{\tau_\Phi^2}}, \\
\sigma_\Phi^2 &= \eta \cdot E\left[\Phi_j^p\right].
\end{aligned}$$

Dôkaz
Stačí si uvedomiť, že za podmienky, že $\mathbf{\Phi} = (\Phi_1, \Phi_2, \ldots \Phi_J)$ je dané, náhodné

veličiny $X_{ij}^{(1)}$ a n_{ij}, $i = 1, 2, \ldots I$, $j = 1, 2 \ldots J$ splňuju predpoklady Bühlmann Straubovho modelu definovaného Tvrdením 4.2.1, kde

$$\begin{aligned} E\left[X_{ij}^{(1)}|\Theta_{ij}\right] &= \mu\left(\Psi_i\right) = \Psi_i, \\ Var\left(X_{ij}^{(1)}|\Theta_{ij}\right) &= \frac{\sigma^2\left(\Psi_i\right)}{n_{ij}^{(1)}} = \frac{\eta\Psi_i^p}{n_{ij}^{(1)}}. \end{aligned} \tag{4.28}$$

Teda $\Psi_i^*\left(\mathbf{\Phi}\right)$ je kredibilitným odhadom parametru Ψ_i na základe $X_{ij}^{(1)}$. Odhad parametru $\Phi_j, j = 1, 2 \ldots J$ možno dokázať analogicky.

Veta 4.2.3.

1. *Odhady parametrov $\Psi_i, \Phi_j, i = 1 \ldots I, j = 1 \ldots J$ založené na kredibilitných odhadoch v modeloch s multiplikatívnou štruktúrou (označíme s horným indexom CM) majú tvar*

$$\Psi_i^{CM} = \Psi_i^*\left(\mathbf{\Phi^{CM}}\right), \tag{4.29}$$

$$\Phi_j^{CM} = \Phi_j^*\left(\mathbf{\Psi^{CM}}\right), \tag{4.30}$$

   kde

$$\begin{aligned} \mathbf{\Psi^{CM}} &= \left(\Psi_1^{CM}, \Psi_2^{CM} \ldots \Psi_I^{CM}\right), \\ \mathbf{\Phi^{CM}} &= \left(\Phi_1^{CM}, \Phi_2^{CM} \ldots \Phi_J^{CM}\right). \end{aligned}$$

2. *Odhad poistného založený na kredibilitnom odhade v modely s multiplikatívnou tarifnou štruktúrou má tvar:*

$$P_{ij}^{CM} = \mu_0 \Psi_i^{CM} \Phi_j^{CM}. \tag{4.31}$$

Odhady vyplývajúce zo vzťahov 4.29 a 4.30 nie sú kredibilitnými odhadmi parametrov ψ_i, ϕ_j, $i = 1, 2 \ldots I$, $j = 1, 2 \ldots J$, lebo nie sú lineárnymi funkciami realizácií náhodných veličín X_{ij}, $i = 1, 2, \ldots I$, $j = 1, 2 \ldots J$, ale sú na kredibilitných odhadoch založené. Riešenie 4.29 a 4.30 možno nájsť použitím iteračných metód.

4.3 Odhady štrukturálnych parametrov

Na to aby sme mohli používať výsledky z predchádzajúcich kapitol, budeme musieť vyčísliť parametre $\sigma^2, \tau_\Phi^2, \tau_\Psi^2$, ktoré nepoznáme. Odhady týchto parametrov si

označíme $\hat{\sigma^2}, \hat{\tau^2_\Psi}, \hat{\tau^2_\Phi}$.
My budeme pri výpočtoch používať nasledujúce odhady:
Pre σ^2

$$\hat{\sigma^2} = \frac{1}{(I-1)(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \left(X_{ij} - \bar{X_{\bullet\bullet}}\right)^2. \tag{4.32}$$

Pre τ^2_Ψ a τ^2_Φ môžeme použiť odhady

$$\begin{aligned} \hat{\tau}^2_\Psi &= max\left(\tilde{\tau}^2_\Psi, 0\right), \\ \hat{\tau}^2_\Phi &= max\left(\tilde{\tau}^2_\Phi, 0\right), \end{aligned} \tag{4.33}$$

kde

$$\begin{aligned} \tilde{\tau}^2_\Psi &= c_1 \left(\frac{I}{I-1} \sum_{i=1}^{I} \frac{n_{i\bullet}}{n_{\bullet\bullet}} \left(\bar{X}_{i\bullet} - \bar{X}_{\bullet\bullet}\right)^2 - \frac{I}{I-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{n_{i\bullet}}{n_{\bullet\bullet}} \left(\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}} \right)^2 \tilde{\tau}^2_\Phi - \frac{I}{n_{\bullet\bullet}} \hat{\sigma}^2 \right), \\ \tilde{\tau}^2_\Phi &= c_2 \left(\frac{J}{J-1} \sum_{j=1}^{J} \frac{n_{\bullet j}}{n_{\bullet\bullet}} \left(\bar{X}_{\bullet j} - \bar{X}_{\bullet\bullet}\right)^2 - \frac{J}{J-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{n_{\bullet j}}{n_{\bullet\bullet}} \left(\frac{n_{ij}}{n_{\bullet j}} - \frac{n_{i\bullet}}{n_{\bullet\bullet}} \right)^2 \tilde{\tau^2_\Psi} - \frac{J}{n_{\bullet\bullet}} \hat{\sigma}^2 \right), \end{aligned} \tag{4.34}$$

a

$$\begin{aligned} c_1 &= \frac{I-1}{I} \left(\sum_i \frac{n_{i\bullet}}{n_{\bullet\bullet}} \left(1 - \frac{n_{i\bullet}}{n_{\bullet\bullet}} \right) \right)^{-1}, \\ c_2 &= \frac{J-1}{J} \left(\sum_j \frac{n_{\bullet j}}{n_{\bullet\bullet}} \left(1 - \frac{n_{\bullet j}}{n_{\bullet\bullet}} \right) \right)^{-1}. \end{aligned}$$

V prípade $\hat{\sigma^2}, \tilde{\tau^2_\Phi}, \tilde{\tau^2_\Psi}$ ide o nestranné odhady parametrov. K vyčísleniu odhadov parametrov $\tilde{\tau^2_\Phi}, \tilde{\tau^2_\Psi}$ je však nutné vyriešiť sústavu rovníc 4.34, ktorá sa zvyčajne rieši iteračne. Práve kvôli tejto komplikácií sa zanedbáva prostredný člen, o ktorom sa predpokladá, že je veľmi malý. Tým získavame nasledujúce odhady štrukturálnych parametrov τ^2_Ψ, τ^2_Φ, ktoré už síce nie sú nestranné, ale ich výhodou je práve ich explicitné vyjadrenie.
Teda

$$\begin{aligned} \tilde{\tau}^2_\Psi &= c_1 \left(\frac{I}{I-1} \sum_{i=1}^{I} \frac{n_{i\bullet}}{n_{\bullet\bullet}} \left(\bar{X}_{i\bullet} - \bar{X}_{\bullet\bullet}\right)^2 - \frac{I}{n_{\bullet\bullet}} \hat{\sigma}^2 \right), \\ \tilde{\tau}^2_\Phi &= c_2 \left(\frac{J}{J-1} \sum_{j=1}^{J} \frac{n_{\bullet j}}{n_{\bullet\bullet}} \left(\bar{X}_{\bullet j} - \bar{X}_{\bullet\bullet}\right)^2 - \frac{J}{n_{\bullet\bullet}} \hat{\sigma}^2 \right). \end{aligned} \tag{4.35}$$

Kapitola 5

Príklad

V tejto kapitole budeme aplikovať metódy, ktoré boli predstavené v predchádzajúcich kapitolách na konkrétnych dátach. Všetky metódy budeme aplikovať na *multiplikatívnu tarifnú štruktúru.*

Ako bolo spomenuté, dve z možných interpretácií μ_0 v 1.3 sú interpretácie μ_0 ako priemerného poistného alebo ako poistného v referenčnej triede, ktorú si dopredu zvolíme.

V prípade interpretácie μ_0 ako priemerného poistného musia odhady parametrov ψ_i, $i = 1, 2 \ldots I$ a ϕ_j, $j = 1, 2 \ldots J$ spĺňať rovnosť 2.12. Odhadnuté parametre porovnávajú rizikovosť príslušnej zmluvy a zmluvy s priemerným poistným z pohľadu jednotlivých tarifných premenných.

Pre prípad interpretácie μ_0 ako poistného v referenčnej triede budeme za referenčnú triedu v celej tejto kapitole považovať tarifnú triedu $(1, 1)$. Parametre v referenčnej triede musia spĺňať nasledovné.

$$\begin{aligned} \hat{P_{1,1}^{M}} &= \tilde{\mu_0}, \\ \tilde{\psi_1} &= 1, \\ \tilde{\phi_1} &= 1, \end{aligned}$$

Tým získame možnosť interpretovať $\tilde{\psi}_i$ ako vyjadrenie rizikovosti zmlúv v tarifnej triede (i, j) oproti zmluvám v triede referenčnej z pohľadu tarifnej premennej A a $\tilde{\phi}_j$ ako vyjadrenie rizikovosti zmluv tarifnej triedy (i, j) oproti zmluvám v triede referenčnej z pohľadu tarifne premennej B.

5.1 Vstupy

Našou úlohou bude modelovať priemernú výšku škody v nasledujúcom období na základe údajov o výške priemernej škody v období minulom. Miera expozície je počet škôd v minulom období. Použijeme zjednodušenú predstavu o dvoch tarifných premenných, označíme A a B. Tarifná premenná A môže nadobúdať 4 hodnoty a tarifná premenná B môže nadobúdať 8 hodnôt. Tým dostaneme počet tarifných tried 32. Data, s ktorými budeme pracovať sú dané nasledujúcimi dvoma tabuľkami.

Tabuľka 5.1: Priemerná veľkosť škody

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	3883	3540	3324	3206	3062	3117	3182	3197
2	3379	3501	3769	3426	3818	4335	2800	2657
3	3923	3575	3265	2484	2896	2764	9169	3553
4	3966	3652	3769	4830	3939	2780	3103	1974

Tabuľka 5.2: Počet škôd

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	2161	10 650	6239	2746	1870	1478	1306	974
2	251	864	501	228	209	103	48	31
3	184	644	261	64	23	15	3	2
4	427	1427	683	105	56	24	9	10

5.2 Základné modely

V tomto odstavci budeme aplikovať metódu Baileyho - Simonovu (MBS), váženú metódu najmenších štvorcov (VMNS) a metódu marginálnych súčtov (MMS). Podrobnejšie sa budeme zaoberať práve MMS, kde si predvedieme interpretáciu μ_0 ako priemerného poistného i ako poistného v referenčnej triede. Vo zvyšných dvoch metódach si uvedieme iba výsledky pre porovnanie jednotlivých metód.

Metóda marginálnych súčtov

1. **Nech μ_0 je priemerné poistné**
   K výpočtu odhadov parametrov použijeme Vetu 2.2.2. Výsledky budú nasledovné.

$$\tilde{\mu_0} = 3\,445.04, \tag{5.1}$$

Tabuľka 5.3: Odhady parametrov ψ_i, $i = 1, 2, \ldots 4$ získané metódou MMS

$\tilde{\psi}_1$	$\tilde{\psi}_2$	$\tilde{\psi}_3$	$\tilde{\psi}_4$
0.990462	1.03742	0.990008	1.06724

Tabuľka 5.4: Odhady parametrov ϕ_j, $j = 1, 2, \ldots 8$ získané metódou MMS

$\tilde{\phi}_1$	$\tilde{\phi}_2$	$\tilde{\phi}_3$	$\tilde{\phi}_4$	$\tilde{\phi}_5$	$\tilde{\phi}_6$	$\tilde{\phi}_7$	$\tilde{\phi}_8$
1.11333	1.02921	0.983861	0.950125	0.91892	0.9299	0.930285	0.92675

V prípade tarifnej triedy číslo $(2, 3)$ máme nasledujúce výsledky.

$$\begin{aligned} \tilde{\mu_0} &= 3\,445.04, \\ \tilde{\psi_2} &= 1.03742, \\ \tilde{\phi_3} &= 0.983861, \\ \tilde{P}^M_{2,3} &= 3\,516.26. \end{aligned}$$

To možno interpretovať nasledovne. Očakávaná priemerná škoda (poistné) je 3 445.04. Z pohľadu tarifnej premennej A je zmluva v $(2, 3)$ rizikovejšia ako zmluva s priemerným poistným, $\left(\tilde{\psi_2} > 1\right)$, a naopak z pohľadu tarifnej premennej B sa nejaví ako výrazne riziková $\left(\tilde{\phi_3} < 1\right)$. Inak povedané, vplyvom tarifnej premennej A sa poistné v $(2, 3)$ zvýši a vplyvom tarifnej premennej B sa zníži, čo má za následok, že v budúcom období očakávame v rámci tejto tarifnej triedy poistné 3277.79.

Výsledné poistné v jednotlivých tarifných triedach je uvedené v Tabuľke 5.5.

Tabuľka 5.5: Výsledné poistné získané metódou MMS

	1	2	3	4	5	6	7	8
1	3798.87	3511.84	3357.11	3242.	3135.52	3172.99	3174.3	3162.24
2	3978.97	3678.32	3516.26	3395.69	3284.16	3323.41	3324.78	3312.15
3	3797.13	3510.23	3355.57	3240.51	3134.08	3171.53	3172.84	3160.79
4	4093.33	3784.05	3617.33	3493.29	3378.56	3418.93	3420.35	3407.35

2. **Nech μ_0 je poistné v referenčnej triede** $(1, 1)$.
   Po vyriešení sústavy rovníc 2.15 dostávame nasledujúce výsledky.
   Poistné v referenčnej triede je

$$\tilde{\mu_0} = 3\,798.87. \tag{5.2}$$

Tabuľka 5.6: Odhady parametrov ψ_i, $i = 1, 2, \ldots 4$ vzhľadom k referenčnej triede

$\tilde{\psi}_1$	$\tilde{\psi}_2$	$\tilde{\psi}_3$	$\tilde{\psi}_4$
1.	1.04741	0.999542	1.07751

Tabuľka 5.7: Odhady parametrov ϕ_j, $j = 1, 2, \ldots 8$ vzhľadom k referenčnej triede

$\tilde{\phi}_1$	$\tilde{\phi}_2$	$\tilde{\phi}_3$	$\tilde{\phi}_4$	$\tilde{\phi}_5$	$\tilde{\phi}_6$	$\tilde{\phi}_7$	$\tilde{\phi}_8$
1.	0.924442	0.883713	0.853411	0.825382	0.835244	0.83559	0.832415

Pre tarifnú triedu $(2, 3)$ máme nasledujúce výsledky.

$$\begin{aligned} \tilde{\mu_0} &= 3\,798.87, \\ \tilde{\psi_2} &= 1.04741, \\ \tilde{\phi_3} &= 0.883713, \\ \tilde{P^M_{2,3}} &= 3\,516.26. \end{aligned}$$

Z výsledkov vyplývy, že z pohľadu tarifnej premennej A je zmluva v referenčnej triede menej riziková ako zmluva v tarifnej triede $(2, 3)$ a z pohľadu tarifnej premennej B je zmluva v referenčnej triede viac riziková ako zmluva v tarifnej triede $(2, 3)$.

Poistné v jednotlivých tarifných triedach vzhľadom k referenčnej triede bude nasledovné.

Tabuľka 5.8: Poistné vzhľadom k referenčnej triede získané metódou MMS

	1	2	3	4	5	6	7	8
1	1.	0.924442	0.883713	0.853411	0.825382	0.835244	0.83559	0.832415
2	1.04741	0.968268	0.925607	0.893868	0.864511	0.874841	0.875203	0.871877
3	0.999542	0.924019	0.883308	0.85302	0.825004	0.834862	0.835207	0.832034
4	1.07751	0.996098	0.952212	0.919561	0.889359	0.899986	0.900359	0.896938

Po vynásobení prvkov predchádzajúcej tabuľky poistným v referenčnej triede dostaneme výsledné poistné.

Tabuľka 5.9: Výsledné poistné získané metódou MMS

	1	2	3	4	5	6	7	8
1	3798.87	3511.84	3357.11	3242.	3135.52	3172.99	3174.3	3162.24
2	3978.96	3678.32	3516.26	3395.69	3284.16	3323.41	3324.78	3312.15
3	3797.13	3510.23	3355.57	3240.51	3134.08	3171.53	3172.85	3160.79
4	4093.33	3784.05	3617.33	3493.29	3378.56	3418.93	3420.35	3407.35

V nasledujúcich dvoch tabuľkách sú uvedené výsledky i zvyšných dvoch metódach pri interpretácií μ_0 ako poistného v referenčnej triede.

Tabuľka 5.10: Odhady neznámych parametrov získané metódami MBS, VMNS, MMS

	MBS	**VMNS**	**MMS**
$\tilde{\mu}_0$	3797.99	3798.39	3798.87
$\tilde{\psi}_1$	1.	1.	1.
$\tilde{\psi}_2$	1.05372	1.04205	1.04741
$\tilde{\psi}_3$	1.00591	1.00059	0.999542
$\tilde{\psi}_4$	1.0823	1.07505	1.07751
$\tilde{\phi}_1$	1.	1.	1.
$\tilde{\phi}_2$	0.923747	0.924699	0.924442
$\tilde{\phi}_3$	0.883233	0.884754	0.883713
$\tilde{\phi}_4$	0.855684	0.854812	0.853411
$\tilde{\phi}_5$	0.826448	0.826824	0.825382
$\tilde{\phi}_6$	0.837873	0.83619	0.835244
$\tilde{\phi}_7$	0.83922	0.835631	0.83559
$\tilde{\phi}_8$	0.833798	0.832185	0.832415

Tabuľka 5.11: Poistné vzhľadom k referenčnej triede získané metódami MBS, VMNS, MMS

i	j	MBS	VMNS	MMS
1	1	1.	1.	1.
1	2	0.923747	0.924699	0.924442
1	3	0.883233	0.884754	0.883713
1	4	0.855684	0.854812	0.853411
1	5	0.826448	0.826824	0.825382
1	6	0.837873	0.83619	0.835244
1	7	0.83922	0.835631	0.83559
1	8	0.833798	0.832185	0.832415
2	1	1.05372	1.04205	1.04741
2	2	0.973375	0.963581	0.968268
2	3	0.930684	0.921957	0.925607
2	4	0.901655	0.890756	0.893868
2	5	0.870848	0.861591	0.864511
2	6	0.882887	0.871351	0.874841
2	7	0.884307	0.870767	0.875203
2	8	0.878592	0.867177	0.871877
3	1	1.00591	1.00059	0.999542
3	2	0.929208	0.925243	0.924019
3	3	0.888454	0.885274	0.883308
3	4	0.860742	0.855315	0.85302
3	5	0.831334	0.82731	0.825004
3	6	0.842826	0.836682	0.834862
3	7	0.844181	0.836122	0.835207
3	8	0.838726	0.832675	0.832034
4	1	1.0823	1.07505	1.07751
4	2	0.999776	0.994096	0.996098
4	3	0.955927	0.951154	0.952212
4	4	0.926111	0.918965	0.919561
4	5	0.894469	0.888876	0.889359
4	6	0.906834	0.898945	0.899986
4	7	0.908292	0.898343	0.900359
4	8	0.902423	0.89464	0.896938

5.3 Zobecnené lineárne modely

Ako bolo v kapitole o GLM avízované, budeme pri modelovaní veľkosti škôd uvažovať Gamma rozdelenie. Hlavné charakteristiky tohoto rozdelenia sú uvedené v 3.16.
S ohľadom na Tabuľku 3.1 dostaneme nasledujúce odhady neznámych parametrov.

Poistné v referenčnej triede je

$$\tilde{\mu_0} = 3\ 800.08047 \tag{5.3}$$

Tabuľka 5.12: Odhady parametrov ψ_i, $i = 1, 2 \ldots 4$ získané metódou GLM

$\tilde{\psi}_1$	$\tilde{\psi}_2$	$\tilde{\psi}_3$	$\tilde{\psi}_4$
1.000000000	1.052846268	0.998386527	1.080089495

Tabuľka 5.13: Odhady parametrov ϕ_j, $j = 1, 2 \ldots 8$ získané metódou GLM

$\tilde{\psi}_1$	$\tilde{\psi}_2$	$\tilde{\psi}_3$	$\tilde{\psi}_4$	$\tilde{\psi}_5$	$\tilde{\psi}_6$	$\tilde{\psi}_7$	$\tilde{\psi}_8$
1.000000000	0.924085082	0.882498617	0.851863720	0.823696394	0.834016275	0.835451870	0.832541196

Tabuľka 5.14: Poistné vzhľadom k referenčnej triede odhadnuté metódou GLM

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	1.00000000	0.92408508	0.88249862	0.85186372	0.82369639	0.83401627	0.83545187	0.83254120
2	1.05284627	0.97291953	0.92913538	0.89688154	0.86722567	0.87809092	0.87960238	0.87653789
3	0.99838653	0.92259410	0.88107473	0.85048926	0.82236738	0.83267061	0.83410389	0.83119791
4	1.08008949	0.99809459	0.95317749	0.92008905	0.88966582	0.90081222	0.90236279	0.89921900

Výsledné poistné v jednotlivých tarifných triedach bude

Tabuľka 5.15: Výsledné poistné získané metódou GLM

	1	2	3	4	5	6	7	8
1	3800.08047	3511.59767	3353.56576	3237.15069	3130.11258	3169.32896	3174.78434	3163.72354
2	4000.90054	3697.17251	3530.78920	3408.22202	3295.52735	3336.81617	3342.55984	3330.91452
3	3793.94914	3505.93180	3348.15487	3231.92763	3125.06223	3164.21533	3169.66191	3158.61896
4	4104.42700	3792.83976	3622.15115	3496.41245	3380.80172	3423.15891	3429.05121	3417.10456

To ako veľmi sa realizácie náhodných veličín a odhadnuté hodnoty líšia budeme skúmať pomocou Pearsonovej χ^2 štatistiky. Jej hodnota je v tomto prípade

$$\chi^2 = 4,2898 \tag{5.4}$$

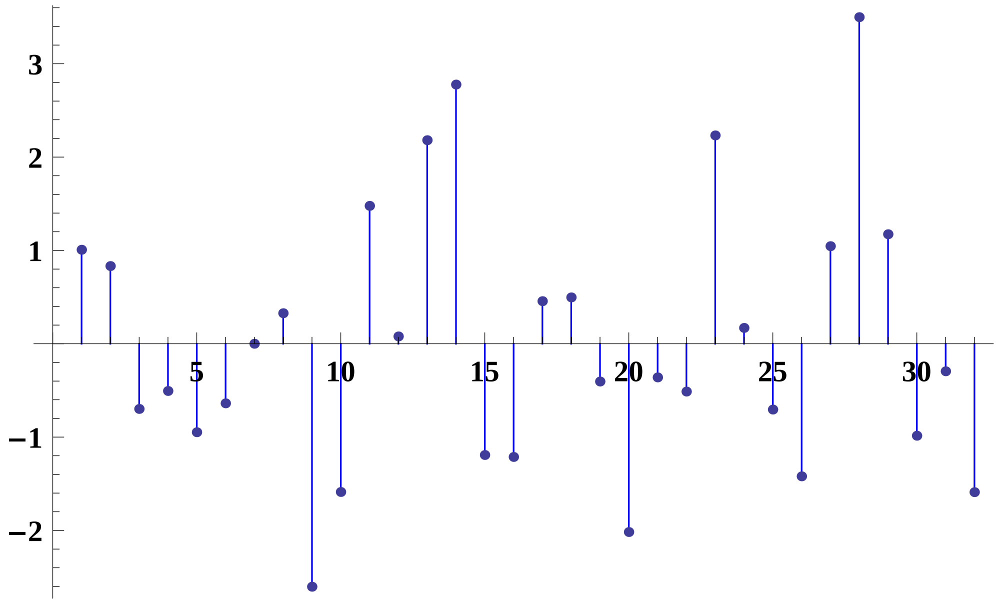
Vzhľadom k Tabuľke 5.16 možno prehlásiť, že rozdiel medzi realizáciami náhodných veličína odhadu ich stredných hodnôt je na príslušných hladiných spoľahlivosti tolerovateľný. Napríklad na hladine spoľahlivosti $0,05$ je $\chi^2_{19}\,(0,05) > \chi^2$.

Tabuľka 5.16: Kvantily χ^2_{19}

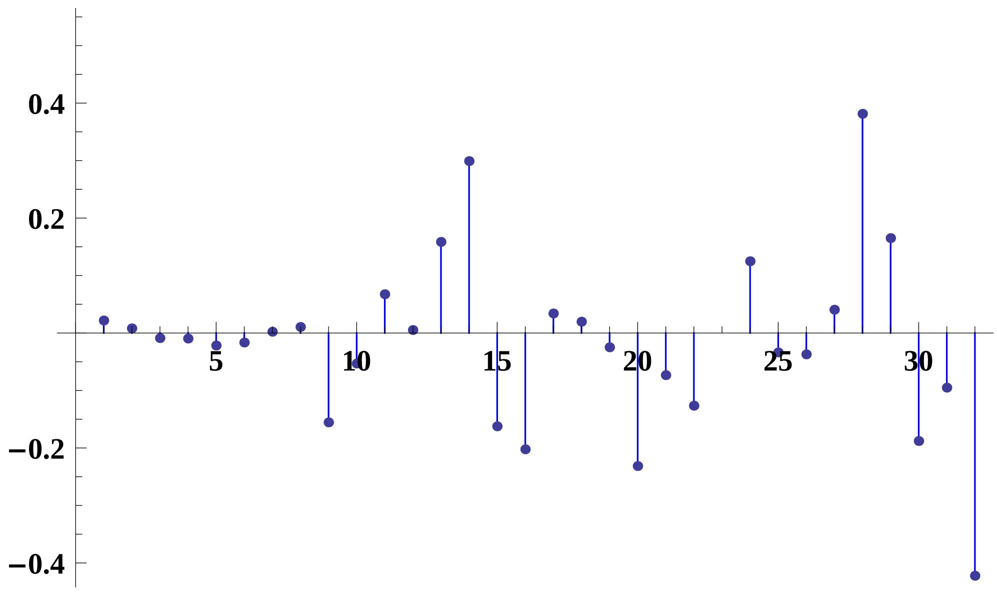
0.005	**0.01**	**0.025**	**0.05**	**0.01**	**0.9**	**0.925**	**0.95**	**0.975**	**0.99**	**0.995**
6.84392	7.6327	8.90651	10.117	7.6327	27.2036	28.4581	30.1435	32.8523	36.1908	38.5821

Pre úplnosť treba uviesť ako sa budú správať reziduá a to ako Pearsonove, definované v 3.39 , tak i deviančné, definované v 3.37.

Obr. 5.1: Deviančné reziduá



Obr. 5.2: Pearsonove reziduá



5.4 Kredibilitné odhady

K výpočtu kredibilitných odhadov neznámych parametrov je najskôr potrebné odhadnúť príslušné štrukturálne parametre. Jednotlivé vzťahy pre tieto odhady boli uvedené v Kapitole 4.3., pričom pre odhady štrukturálnych parametrov τ_Ψ^2 a τ_Φ^2 použijeme $\tilde{\tau}_\Psi^2, \tilde{\tau}_\Phi^2$ uvedené vzťahmi 4.35. Odhady budú nasledovné.

Tabuľka 5.17: Odhady štrukturálnych parametrov

$\hat{\mu}_0$	$\hat{\sigma}^2$	$\tilde{\tau}_\psi^2$	$\tilde{\tau}_\phi^2$
3445.04	1.08506×10^8	5202.33	19 287.4

Keďže ide o Gamma rozdelenie, tak $p = 2$. Za použitia vzťahov predstavených v Kapitole 4 možno dospieť k nasledujúcim výsledkom.

Pomocou Vety 4.2.3 a interpretácie μ_0 ako poistného v referenčnej triede dostaneme nasledujúce výsledky.

Poistné v referenčnej triede je

$$\tilde{\mu_0} = 3\,552.74.$$

Tabuľka 5.18: Kredibilitné odhady parametrov ψ_i, $i = 1, 2, \ldots 4$

$\tilde{\psi}_1$	$\tilde{\psi}_2$	$\tilde{\psi}_3$	$\tilde{\psi}_4$
1.	1.01601	1.0127	1.02163

Tabuľka 5.19: Kredibilitné odhady parametrov ϕ_j, $j = 1, 2, \ldots 8$

$\tilde{\phi}_1$	$\tilde{\phi}_2$	$\tilde{\phi}_3$	$\tilde{\phi}_4$	$\tilde{\phi}_5$	$\tilde{\phi}_6$	$\tilde{\phi}_7$	$\tilde{\phi}_8$
1.	0.984348	0.953522	0.942325	0.937436	0.943319	0.944973	0.947113

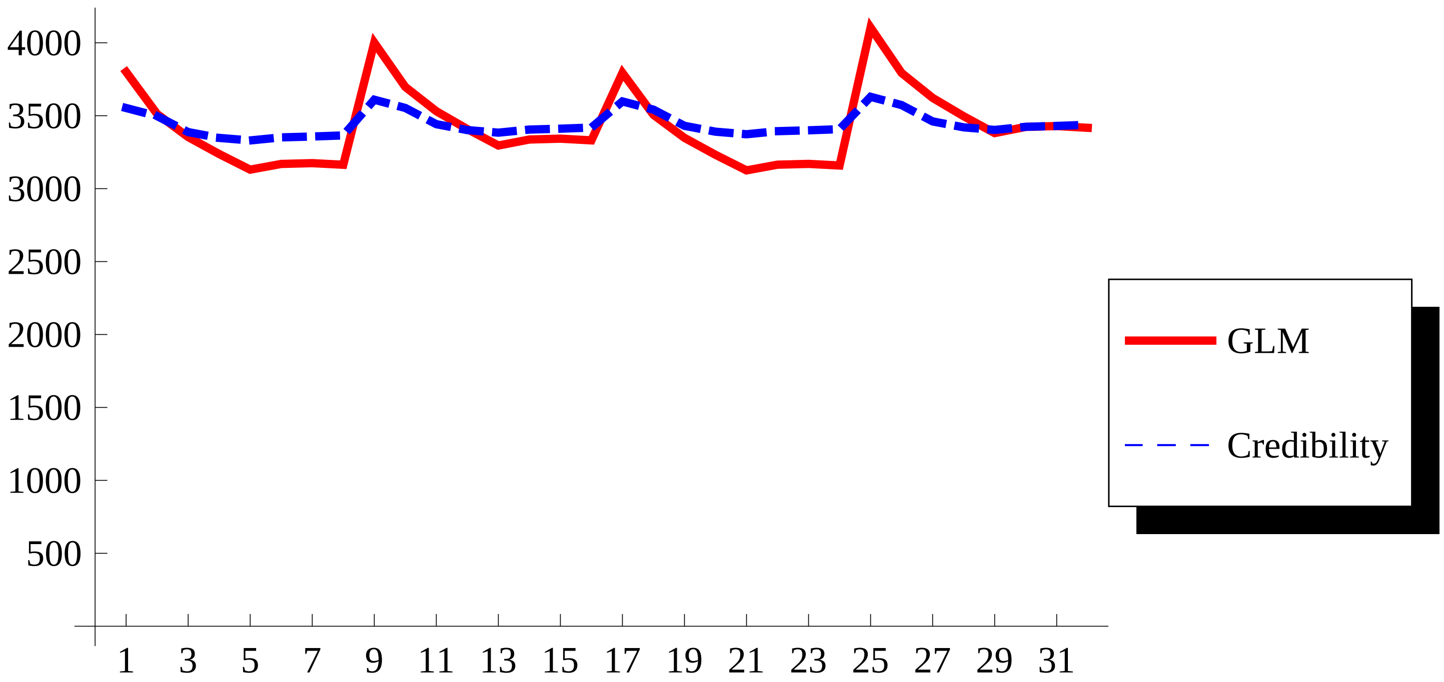
Tabuľka 5.20: Kredibilitné poistné vzhľadom k referenčnej triede

	1	2	3	4	5	6	7	8
1	1.	0.984348	0.953522	0.942325	0.937436	0.943319	0.944973	0.947113
2	1.01601	1.00011	0.968786	0.957409	0.952442	0.958419	0.9601	0.962274
3	1.0127	0.996852	0.965635	0.954295	0.949344	0.955302	0.956977	0.959144
4	1.02163	1.00564	0.974149	0.962709	0.957714	0.963725	0.965415	0.967601

Tabuľka 5.21: Výsledné poistné určené metódou kredibilitných odhadov

	1	2	3	4	5	6	7	8
1	3552.74	3497.13	3387.62	3347.83	3330.46	3351.37	3357.24	3364.85
2	3609.61	3553.11	3441.85	3401.43	3383.78	3405.01	3410.98	3418.71
3	3597.87	3541.56	3430.65	3390.36	3372.77	3393.94	3399.89	3407.59
4	3629.59	3572.78	3460.9	3420.25	3402.51	3423.86	3429.87	3437.64

Jednou z výhod kredibilitných odhadov je i vyhladzovanie poistného medzi jednotlivými triedami. Vyhladzovací efekt možno pozorovať na nasledujúcom obrázku, ktorý zároveň porovnáva metódu GLM s kredibilitnými odhadmi.



Obr. 5.3: Porovnanie kredibilitných odhadov a GLM

Tabuľka 5.22 zhŕňa výsledné poistné vypočítané jednotlivými metódami.

Tabuľka 5.22: Výsledné poistné

i	j	MBS	VMNS	MMS	GLM	KO
1	1	3797.99	3798.39	3798.87	3800.08047	3552.74
1	2	3508.39	3512.36	3511.84	3511.59767	3497.13
1	3	3354.51	3360.64	3357.11	3353.56576	3387.62
1	4	3249.88	3246.91	3242.	3237.15069	3347.83
1	5	3138.85	3140.6	3135.52	3130.11258	3330.46
1	6	3182.24	3176.17	3172.99	3169.32896	3351.37
1	7	3187.35	3174.05	3174.3	3174.78434	3357.24
1	8	3166.76	3160.96	3162.24	3163.72354	3364.85
2	1	4002.04	3958.1	3978.96	4000.90054	3609.61
2	2	3696.87	3660.05	3678.32	3697.17251	3553.11
2	3	3534.73	3501.95	3516.26	3530.78920	3441.85
2	4	3424.48	3383.43	3395.69	3408.22202	3401.43
2	5	3307.48	3272.65	3284.16	3295.52735	3383.78
2	6	3353.2	3309.73	3323.41	3336.81617	3405.01
2	7	3358.59	3307.51	3324.78	3342.55984	3410.98
2	8	3336.89	3293.87	3312.15	3330.91452	3418.71
3	1	3820.44	3800.62	3797.13	3793.94914	3597.87
3	2	3529.12	3514.43	3510.23	3505.93180	3541.56
3	3	3374.34	3362.61	3355.57	3348.15487	3430.65
3	4	3269.09	3248.82	3240.51	3231.92763	3390.36
3	5	3157.4	3142.44	3134.08	3125.06223	3372.77
3	6	3201.05	3178.04	3171.53	3164.21533	3393.94
3	7	3206.19	3175.91	3172.85	3169.66191	3399.89
3	8	3185.48	3162.82	3160.79	3158.61896	3407.59
4	1	4110.59	4083.45	4093.33	4104.42700	3629.59
4	2	3797.14	3775.96	3784.05	3792.83976	3572.78
4	3	3630.6	3612.85	3617.33	3622.15115	3460.9
4	4	3517.36	3490.58	3493.29	3496.41245	3420.25
4	5	3397.19	3376.29	3378.56	3380.80172	3402.51
4	6	3444.15	3414.54	3418.93	3423.15891	3423.86
4	7	3449.69	3412.25	3420.35	3429.05121	3429.87
4	8	3427.4	3398.19	3407.35	3417.10456	3437.64

Poznámka 10.
Pri všetkých výpočtoch bol použitý matematický software Mathematica 7. Zdrojový kód nájdete na priloženom CD.

Dodatok A

Vybrané rozdelenia z exponenciálnej rodiny rozdelení

Table 2.1 *Characteristics of some common univariate distributions in the exponential family*†

	Normal	*Poisson*	*Binomial*	*Gamma*	*Inverse Gaussian*
Notation	$N(\mu, \sigma^2)$	$P(\mu)$	$B(m, \pi)/m$	$G(\mu, \nu)$	$IG(\mu, \sigma^2)$
Range of y	$(-\infty, \infty)$	$0(1)\infty$	$\frac{0(1)m}{m}$	$(0, \infty)$	$(0, \infty)$
Dispersion parameter: ϕ	$\phi = \sigma^2$	1	$1/m$	$\phi = \nu^{-1}$	$\phi = \sigma^2$
Cumulant function: $b(\theta)$	$\theta^2/2$	$\exp(\theta)$	$\log(1+e^\theta)$	$-\log(-\theta)$	$-(-2\theta)^{1/2}$
$c(y;\phi)$	$-\frac{1}{2}\left(\frac{y^2}{\phi} + \log(2\pi\phi)\right)$	$-\log y!$	$\log\binom{m}{my}$	$\nu\log(\nu y) - \log y$ $- \log\Gamma(\nu)$	$-\frac{1}{2}\left\{\log(2\pi\phi y^3) + \frac{1}{\phi y}\right\}$
$\mu(\theta) = E(Y;\theta)$	θ	$\exp(\theta)$	$e^\theta/(1+e^\theta)$	$-1/\theta$	$(-2\theta)^{-1/2}$
Canonical link: $\theta(\mu)$	identity	log	logit	reciprocal	$1/\mu^2$
Variance function: $V(\mu)$	1	μ	$\mu(1-\mu)$	μ^2	μ^3

†The mean-value parameter is denoted by μ, or by π for the binomial distribution.
The parameterization of the gamma distribution is such that its variance is μ^2/ν.
The canonical parameter, denoted by θ, is defined by (2.4). The relationship between μ and θ is given in lines 6 and 7 of the Table.

Literatúra

[1] Bühlmann, H.; Gisler, A.: *A course in credibility theory and its applications* Springer - Verlag Berlin Heidelberg 2005

[2] Cipra, T.: *Pojistná matematika - teorie a praxe* Ekopress s.r.o., 2006

[3] Gisler, A.; Muller, P.: *Credibility for additive and multiplikative models*, ASTIN Colloquium, 2007

[4] Holler, K.D; Sommer, D.;Trahair, G. : *Something old, something new in classification ratemaking with a novel use of GLMs for credit insurance*, CAS Forum, Winter 1999, CAS 31 - 84

[5] Mandl, P.; Mazurová, L.; *Matematické modelování v neživotním pojištení* Matfyzpress, 1999

[6] McCullagh, P.; Nelder, J.A.; *Generalized linear models* Second edition, Chapman Hall, 1989

[7] Van Eeghen, J.; Greup, E.K.;Nijssen, J.A.: *Rate Making*, Surveys of actuarial studies No. 2, Nationale - Nederlanden, Rotterdam, 1983

[8] Šimurda, M.: *Zobecněný lineární model: od základů k aplikaci v analýze chování pojistného kmene a škodních událostí* Seminář z aktuářských věd 2007/08, Matfyzpress, 2008